STRUCTURE AND REGULATION OF

YEAST GLYCOGEN SYNTHASE

Sulochanadevi Baskaran

Submitted to the faculty of the University Graduate School
in partial fulfillment of the requirements
for the degree
Doctor of Philosophy
in the Department of Biochemistry and Molecular Biology
Indiana University

August 2010

Accepted by the Faculty of Indiana University, in partial
fulfillment of the requirements for the degree of Doctor of Philosophy.

Thomas D. Hurley, Ph.D., Chair

Anna A. DePaoli-Roach, Ph.D.

Doctoral Committee

Millie M. Georgiadis, Ph.D.

Peter J. Roach, Ph.D.

May 19, 2010

William J. Sullivan Jr., Ph.D.

Dedication

I dedicate this work to my family, friends and mentors, for their support, encouragement, inspiration and guidance.

Acknowledgements

The work presented here is a result of the collaborative effort of many people and it is my pleasure to thank everyone who made this possible.

First and foremost, I am indebted to my advisor, Dr. Thomas Hurley, for his invaluable scientific guidance, enthusiastic supervision and encouragement throughout my thesis research work. I am very grateful for his support, patience and his efforts in explaining and teaching me the concepts of macromolecular crystallography. I will always be thankful to him for teaching me the most important factor that is required for scientific research - perseverance. I would like to express my sincere gratitude to my committee members Drs. Peter Roach, Anna DePaoli-Roach, Millie Georgiadis and William Sullivan for their time, support, guidance and helpful suggestions.

I am thankful to lab colleagues Dr. Samantha Perez-Miller, Dr. Heather Larson, Dr. Jianzhong Zhou, Dr. Lillian González-Segura, Ram Vanam, Jason Braid, Lili Zang, Bibek Parajuli, Dr. May Khanna and Dr. Ann Kimble-Hill for providing a stimulating and fun environment in the lab. I am especially grateful to Dr. Samantha Perez-Miller and Dr. Heather Larson for their friendship and support. I appreciate the help of my colleagues, Dr. Kristie Goodwin, Sarah Delaplane, Dr. Hongzhen He, Dr. Tsuyoshi Imasaki, Dr. Alexander Skurat, Dr. Jose Irimia-Dominguez, Vinnie Tagliabracci, Sixin Jiang, Cathy Meyer, Dyann Segvich and Chandra Karthik. I would like to extend my thanks to Drs. Yuichiro Takagi, Mark Goebl, Ronald Wek and Zhong-Yin Zhang for their encouragement and to the biochemistry office staff for their timely help and assistance.

I wish to thank the beamline scientist at the SBC-CAT and GM/CA-CAT stations at the Advanced Photon Source, in particular Drs. Stephan Ginell, Norma Duke, Marianne Cuff, Nagarajan Venugopalan and Michael Becker for their scientific assistance. I would like to thank Dr. Lou Messerle for providing the tantalum cluster compound and Dr. Thomas Terwilliger for sending the scripts for partial model phasing.

I am grateful to my teachers - Drs. Bhaskar-Rao, Sairam, Illango and Dharmalingam for encouraging me to pursue a career in scientific research. I would like to specifically thank Drs. Usha, Krishnaswamy and Mohammed Rafi for introducing me to the world of structural biology and macromolecular crystallography.

I would like to thank my friends Dr. Shankar Varadarajan, Sowmya Chandrasekar, Dr. Raji Muthukrishanan, Sirisha Pochareddy, Dr. Judy Rose James and Dr. Aditi Bapat for their emotional support and helping me get through the difficult times. Finally I would like to express my sincere thanks to my family for everything they had provided me all through my life.

Abstract

Sulochanadevi Baskaran

STRUCTURE AND REGULATION OF
YEAST GLYCOGEN SYNTHASE

Glycogen is a major energy reserve in most eukaryotes and its rate of synthesis is controlled by glycogen synthase. The activity of eukaryotic glycogen synthase is regulated by the allosteric activator glucose-6-phosphate, which can overcome the inhibitory effects of phosphorylation. The effects of phosphorylation and glucose-6-phosphate on glycogen synthase are mediated by a cluster of six arginines located within a stretch of 12 amino acids near the C-terminus of the enzyme's polypeptide chain. We studied isoform-2 of yeast glycogen synthase as a model to study the structural and molecular mechanisms that underlie the regulation of the eukaryotic enzymes and our primary tools of investigation were macromolecular X-ray crystallography, site-directed mutagenesis, intein-mediated peptide ligation and enzyme assays. We have solved the tetrameric structure of the yeast enzyme in two different activity states; the resting enzyme and the activated state when complexed with glucose-6-phosphate. Binding of glucose-6-phosphate to glycogen synthase induces large conformational changes that free the active site of the subunits to undergo conformational changes necessary to catalyze the reaction. Further, using site directed mutagenesis and intein-mediated peptide ligation to create specific phosphorylation states of the enzyme we were able to define specific roles for

the arginine residues that mediate the regulatory effects of phosphorylation and glucose-6-phosphate activation. Based on these studies, we propose a three state structural model for the regulation of the enzyme, which relate the observed conformational states to specific activity levels. In addition to these regulatory studies, we have also solved the structure of the enzyme complexed with UDP and with substrate analogs, which provide detailed insight into the catalytic mechanism of the enzyme and the ability of glycogen synthase to remain tightly bound to its substrate glycogen.

Thomas D. Hurley, Ph.D., Chair

.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AMP	Adenosine monophosphate
ATP	Adenosine triphosphate
BE	Branching enzyme
Bis-Tris	Bis (2-hydroxyethyl)iminotris(hydroxymethyl)-methane
BME	β-mercapto ethanol
c-AMP	cyclic adenosine monophosphate
CBD	Chitin binding domain
CDP	Cytidine diphosphate
DBE	Debranching enzyme
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
EDTA	Ethylene diamine tetra aceticacid
FOM	Figure of merit
G6P	Glucose-6-phosphate
GFP	Green fluorescence protein
GK	Glucokinase
GKRP	Glucokinase regulatory protein
GPh	Glycogen phosphorylase
GS	Glycogen synthase
GSD	Glycogen storage disorder
GSK3	Glycogen synthase kinase 3
Gsy2p	Glycogen synthase isoform 2 protein

GT	Glycosyl transferase
IMPACT	Intein mediated purification with an affinity chitin-binding tag
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LB	Luria-Bertani broth
MESNA	Sodium 2-sulfanylethanesulfonate
NCS	Non crystallographic symmetry
PCR	Polymerase chain reaction
PEG	Poly ethylene glycol
PKA	Protein Kinase A
PMSF	Phenyl methyl sulfonyl fluoride
PP1	Protein phosphates 1
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
STRE	Stress response element
Tris	Tris (hydroxymethyl) aminomethane
UDP	Uridine diphosphate

INTRODUCTION

A. Glycogen

1. Structure of glycogen

Glycogen, a branched polymer of glucose serves as one of the major repositories of carbon and energy in eukaryotes. The linear polymerization of glycogen is through α-1,4 glycosidic bonds and branch points are introduced by α-1,6 linkages, on an average for every ten to thirteen residues. Glycogen molecules are spherical in shape, organized in concentric tiers and the structure is an example of biological fractal where any substructure of the particle is representative of the whole structure[1,2]. It is theorized that the matured glycogen molecule contains 12 tiers, with approximately 55,000 glucose residues and a molecular weight on the order of 10^7 daltons. The spherical structure of glycogen gives a homogeneously symmetrical shape and enables the maximal storage of glucose in the minimal volume while exposing the maximal number of terminal glucose residues on the outer surface. The structural organization also provides stability by facilitating formation of the maximum number of hydrogen bonds between the glucose residues within the same polymer. The fractal organization of glycogen facilitates the rapid synthesis and degradation, allowing quick release of the stored fuel and fast recovery upon depletion. In addition to glucose, glycogen also contains minor constituents like glucosamines[3] and

phosphates[4], the later influences the branching characteristics of glycogen and is implicated in Lafora disease[5].

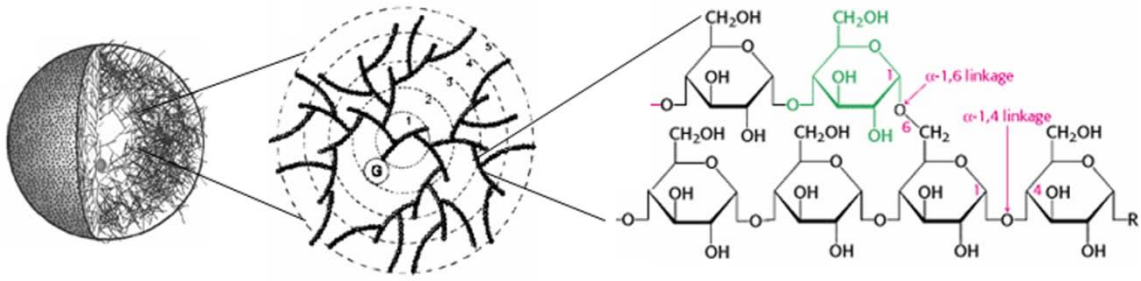

Figure 1. Structutre of glycogen.
The spherical glycogen molecule is proposed to be composed of twelve concentric tires and is formed by linear polymerization via α-1,4 linkages and branch points through α-1,6 linkages. Figure adapted from *Biophys J* . 77, 1327-1332, (1999).

2. Biosynthesis and degradation

The biosynthetic pathway of glycogen synthesis is highly conserved across eukaryotic species (Figure 2). In cells, synthesis from glucose begins with its conversion to UDP-glucose through the sequential action of hexokinase, phosphoglucomutase and UDP-glucose pyrophosphorylase. Glucose polymerization is initiated by glycogenin in an autocatalytic manner[6,7] where glucose residues are transferred from the glucose donor UDP-glucose to a conserved tyrosine residue in the protein, via covalent O-glycosidic linkages[8,9]. Further polymerization of glucose to about 10 residues through α –1,4 glycosidic linkage is required before glycogen synthase (GS) and the branching enzyme (BE) take over. GS catalyzes the linear polymerization of glucose by transferring

glucose residues from an activated sugar donor to the non-reducing 4’ end of the glycogen chain. BE is an amylo (1,4→1,6) – transglycosylase and transfers the terminal chain segment of approximately seven glycosyl residues to the C6 hydroxyl group of a glucose residue on the same or another chain[10].

The two enzymes involved in the degradation of glycogen are glycogen phosphorylase (GPh) and the debranching enzyme (DBE). GPh catalyzes the sequential phosphorolysis of the α –1,4 glycosidic linkages generating glucose-1-phosphate and uses pyridoxal phosphate as the cofactor. When the linear chain is less than five residues from a branch point (limit branch), DBE comes into play. The N-domain α –1,4 transglycosylase activity of DBE transfers the limit branch of glycogen to the 4’ end of another branch facilitating further hydrolysis by GPh[11]. The C-domain α –1,6 glucosidase activity of DBE hydrolyses the α –1,6 linkages at the branch points[11].

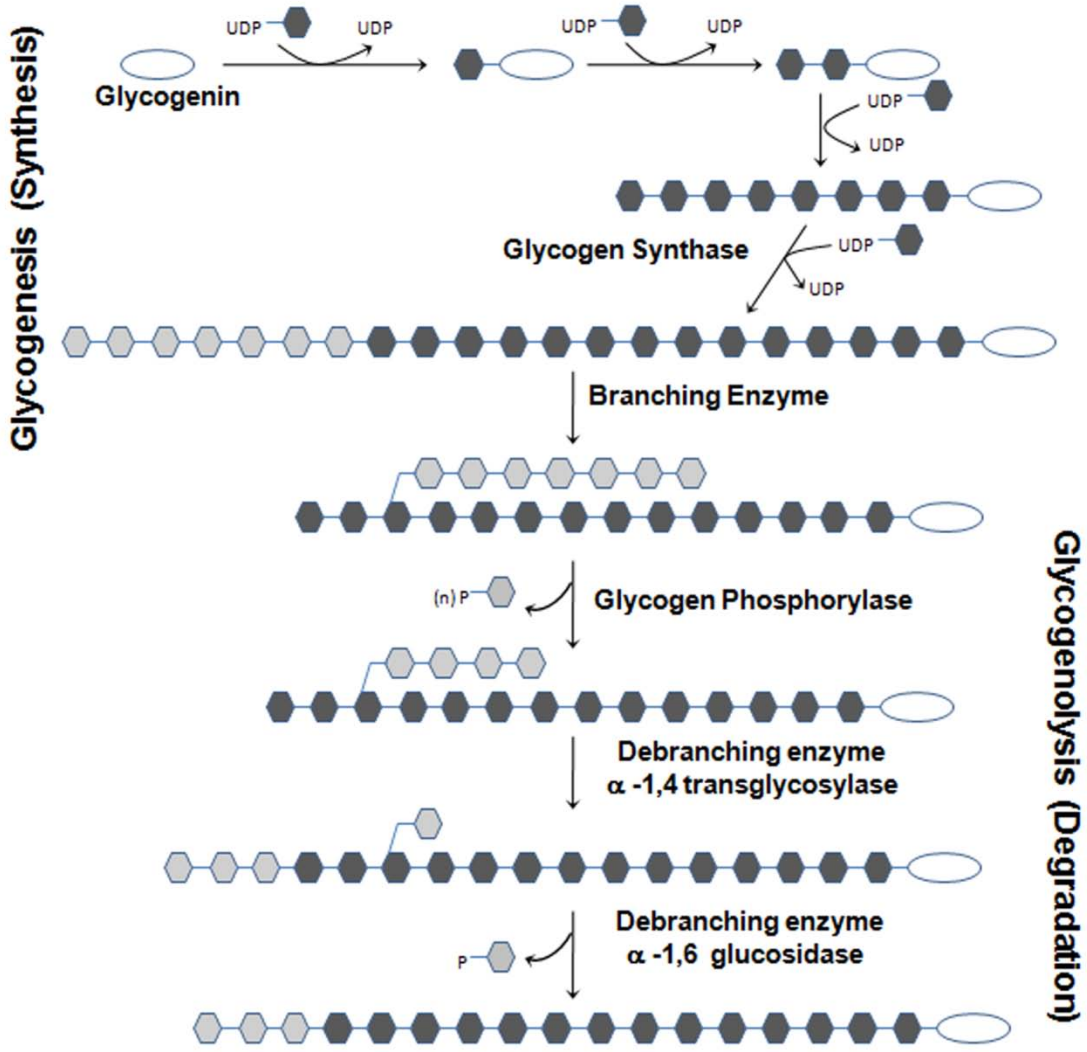

Figure 2. Pathway for the biosynthesis and degradation of glycogen. Synthesis of glycogen polymer involves the activity of the glycogenin, glycogen synthase and branching enzyme. Degradation of the polymer is mediated by glycogen phosphorylase and debranching enzyme.

3. Physiological role of glycogen in mammals

Glycogen serves as the primary reserve of energy in most animals and fungi. Though the biosynthetic pathway of synthesis is highly conserved, the nutritional and hormonal stimuli that regulate the synthesis and degradation of glycogen are different in these organisms. A detailed discussion of the all the regulatory pathways is beyond the scope of this thesis session and a brief overview of the regulation is provided.

In higher eukaryotes including mammals, glycogen is synthesized at times of nutritional abundance[12-14]. The two major tissues or organ systems that serve as the glycogen stores in the higher eukaryotes are skeletal muscle and liver. Other organs like the brain, adipose, kidney and pancreas are also capable of synthesizing glycogen. Insulin stimulated glycogen synthesis accounts for up to 30% in liver and between 30-90% in muscle of postprandial carbohydrate disposal. Depletion of liver and muscle glycogen is observed in type 2 diabetics and impairment of insulin stimulated glycogen synthesis is detectable during the early onset of diabetes and in the pathogenesis of insulin resistant type 2 diabetes[15]. Deficiency in the enzymes involved in glycogen metabolism lead to glycogen storage disease (GSD), which affect the liver, muscle or both tissues.

In the skeletal muscle, glycogen provides energy for muscular contraction in the generation of glucose-6-phosphate for entry into glycolysis as a means for ATP production. The liver glycogen plays a pivotal role in glucose homoeostasis – maintaining circulating blood glucose levels during fasting. The muscle and

liver tissues express different forms of glucose transporters, hexokinases, GPh and GS, and the regulation of glycogen metabolism is slightly different in both these tissues, reflecting their distinct metabolic roles.

i. Skeletal muscle glycogen

The insulin dependent transport of glucose into muscle is mediated by the GLUT4 transporter[16,17]. Upon entry into the cell, glucose is converted to glucose-6-phosphate by the enzyme hexokinase II. In the resting state, glucose-6-phosphate is targeted to either glycolysis for the generation of ATP or the biosynthesis of glycogen. Insulin dependent inhibition of glycogen synthase kinase 3 (GSK3) and dephosphorylation of GS by protein phosphatases promote the synthesis of glycogen[12-14]. Upon initiation of muscular contraction, breakdown of ATP increases cellular AMP levels, which in turn activates glycolysis by stimulating the enzyme phosphofructo kinase.

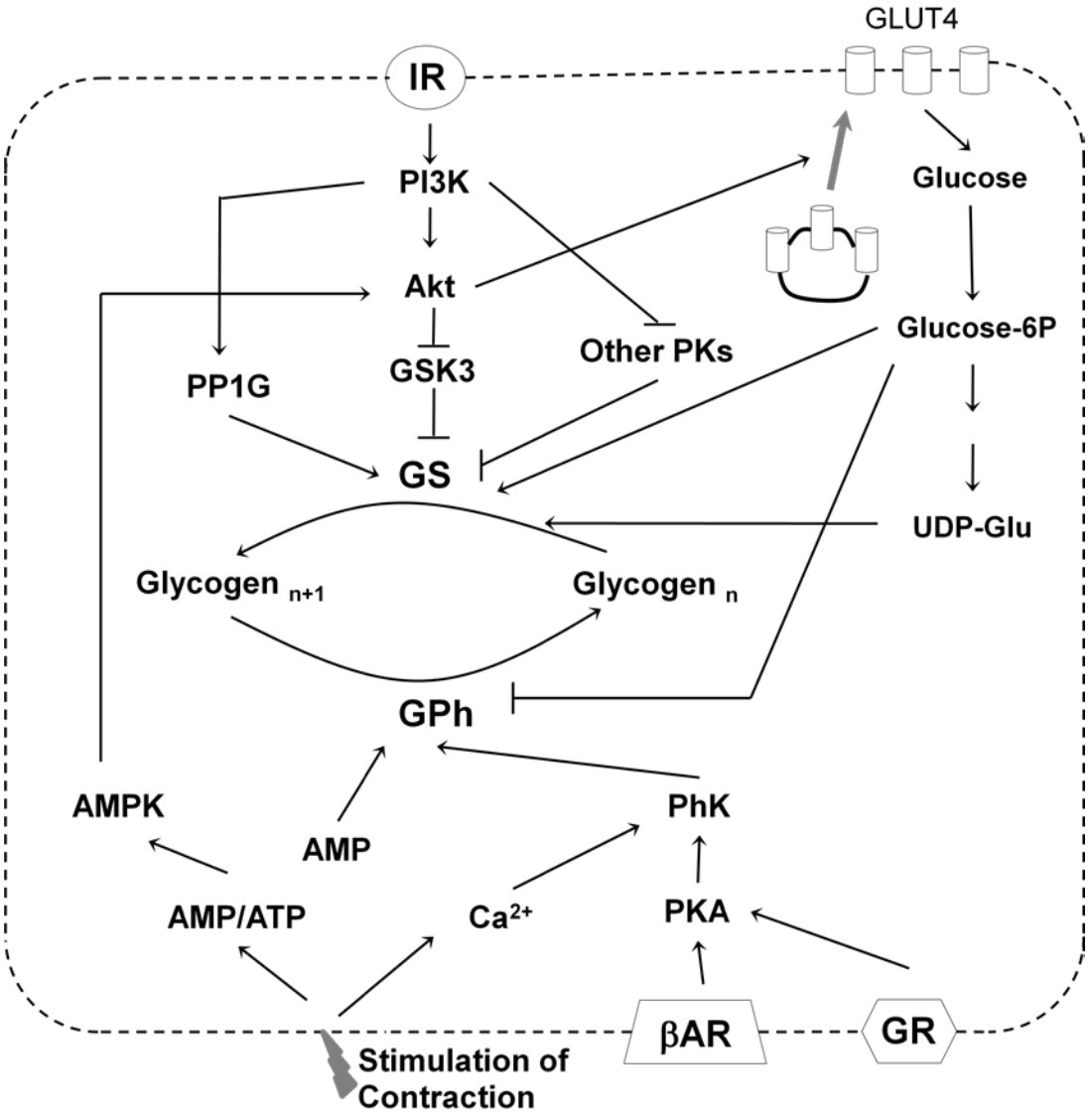

Figure 3. Regulation of glycogen metabolism in skeletal muscle.
Schematic representation of the major signaling pathways regulating glycogen metabolism in the skeletal muscle. IR-Insulin receptor,βAR-βAdrenergic Receptor, GR- Glucagon receptor, PK – Protein Kinase, PI3K – Phosphotidyl inositol kinase, GSK3- Glycogen Synthase Kinase, PP1G- Protein Phosphatase, Gs- Glycogen Synthase, GPh – Glycogen Phosphorylase, AMPK- AMP dependent Protein Kinase, PhK – Phosphorylase Kinase

Release of calcium from the sarcoplasmic reticulum activates phosphorylase kinase, which phosphorylates and activates GPh thus simulating glycogenolysis. Glycogenolysis is also subjected to hormonal activation by epinephrine via protein kinase A (PKA) mediated activation of phosphorylase kinase. As muscular contraction continues, increases in the AMP levels activate AMP kinase, which stimulates glucose uptake. Further the muscle switches fuel utilization and oxidizes fatty acids to produce ATP. Repletion of the glycogen reserve is primarily through the insulin simulated uptake of glucose and GS activation in the fed state.

ii. Liver glycogen

The major glucose transporter in the liver is GLUT2, which is expressed constitutively[18]. Hepatic glucose is phosphorylated by glucokinase (GK) the activity of which is regulated by the glucokinase regulatory protein (GKRP)[19]. GKRP binds to GK and retains it in the nucleus. Binding of glucose or fructose-1P to GK releases GKRP and translocates GK to the cytoplasm. Insulin release promoted by the increased blood glucose levels, stimulates glycogenesis by inactivating glycogen synthase kinase3 (GSK3) and activating protein phosphatases. The decrease in blood glucose by fasting promotes glucagon release from the pancreas, which activates protein kinase A (PKA). PKA activates the phosphorylation cascade involving phosphorylase kinase and glycogen phosphorylase, thus stimulating the breakdown of glycogen.

4. Physiological role of glycogen in yeast

In the budding yeast *Sacccharomyces cerevisiae*, glycogen accounts for 20% of the dry weight of the cells and is one of the two major reserves of carbohydrate, the other being trehalose[20,21]. The amount of glycogen accumulated in the cell increases when the cell enters the stationary phase or upon depletion of essential nutrients like nitrogen and phosphorous in the growth media or by exposing the exponentially growing cells to high temperature, salt, oxidizing agents or ethanol[22].

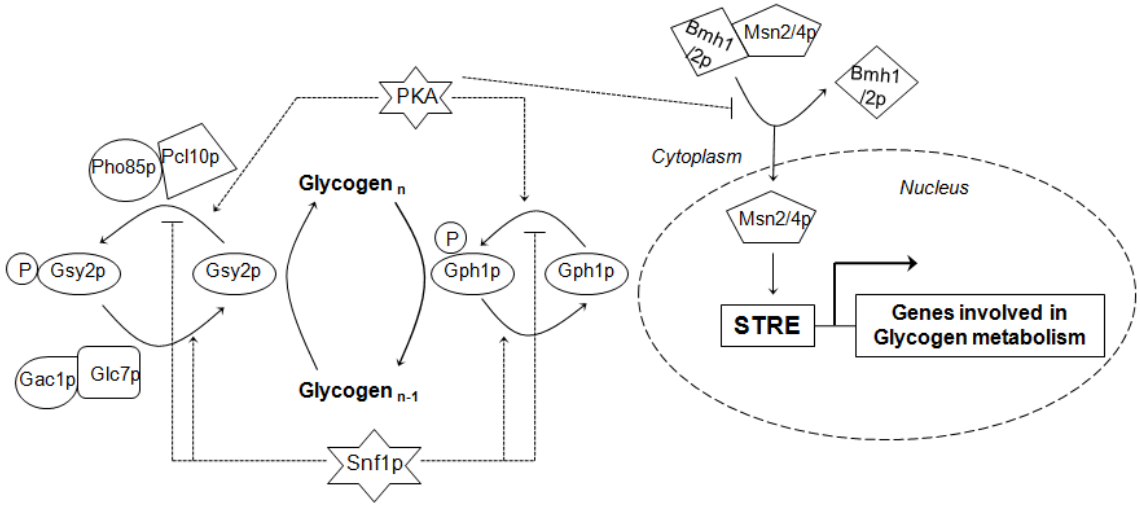

Figure 4. Transcriptional and enzymatic regulation of glycogen metabolism in yeast.
The phosphorylation state of Gsy2p and Gph1p are controlled by PKA and Snf1p in an antagonistic manner. Pho85p in association with Pcl10p phosphorylates and inactivates Gsy2p. The dephosphorylation is catalyzed by the phosphatase Glc7p in association with the targeting subunit Gac1p. The genes involved in glycogen metabolism contain STRE in the promoter and the transcription of the genes is activated by binding of Msn2/4p. PKA negatively controls the gene expression by inhibiting nuclear localization of Msn2p.

Yeast has two different isoforms of GS, of which the nutrionally regulated isoform-2 (GSY2) has shown to be the most important for the accumulation of glycogen in the cells[23]. Unlike the higher eukaryotes where the regulation of glycogen metabolism is primarily through the control of the enzyme activities, in yeast it involves both transcriptional and enzymatic responses. The transcriptional response is dependent on the presence of the cis-element – "stress response element (STRE)" in the promoter of the genes involved in glycogen pathways[24]. Under stress conditions, binding of the trans-activator Msn2p/Msn4p to the STRE causes a 2-3 fold transcriptional activation of these genes. The enzymatic control of glycogen deposition is through the glucose-6-phosphate mediated activation of GS and inactivation of GPh[21] through phosphorylation. Exposure of the starved cells to nutrients activates GPh and inhibits GS resulting in the mobilization of glycogen. The response is biphasic[25] and involves a transient response via the glucose activated c-AMP dependent stimulation of PKA and a sustained response that involves a poorly characterized c-AMP independent fermentable growth medium pathway.

B. Glycogen synthase

1. Glycogen storage disease type-0

GS is one of the rate limiting enzymes in the biosynthetic pathways of glycogen. Deficiency of the human liver glycogen synthase enzyme leads to GSD-0, which is inherited in an autosomal recessive manner. The deficiency causes a marked decrease in the liver glycogen content and is characterized by severe fasting hypoglycemia with associated symptoms including but not limited to lethargy, pallor, nausea and seizures in the morning before food intake. Fasting hypoglycemia is accompanied by hyperketonemia and low blood alanine and lactate levels. GSD-0 has been mapped to the *GYS2* gene located at the chromosomal locus 12p12.2 and sixteen different mutations of *GYS2* have been reported in GSD-0[26-28], which include two splice site variations, four premature stop codons, one deletion mutation and nine missense mutations (Figure 5).

Two recent studies have shown that mutations in the *GYS1* gene could also lead to defects in glycogen storage[29-31]. The presence of a premature stop codon at position 462 of the human muscle GS causes abnormal heart rate and blood pressure after exercise and could lead to hypertropic cardiomyopathy[29]. Muscle biopsy from the patient shows severe lack of glycogen, extenstive mitochondrial proliferation and predominance of oxidative fibers[29]. Similar to the effects in humans, a R309H mutation in horse GYS1 leads to a polysaccharide storage myopathy[30]. The mutation has been reported to increase GS activity with

associated abnormal increase in glycogen accumulation in the skeletal muscles leading to muscle damage with exertion.

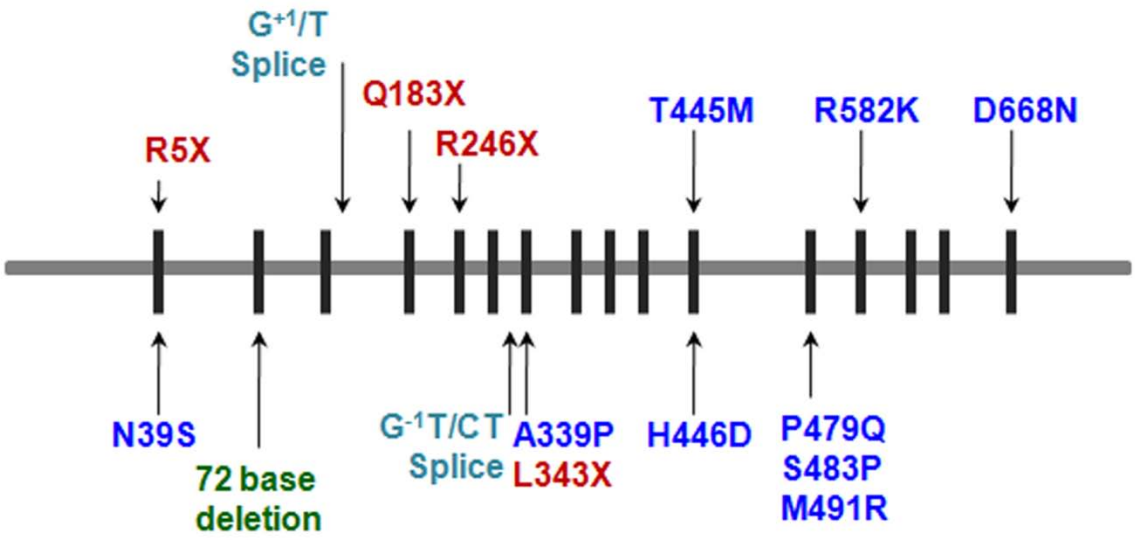

Figure 5. GSD-0 mutation in human GYS2.
Schematic representation of mutations reported in GSD-0 patients. The GYS2 gene is represented as horizontal grey bar and the individual exons vertical black lines. The different types of mutations are color coded.

2. Enzymatic activity of GS

In vitro, GS activity is determined by measuring the amount of ^{14}C-glucose transferred from UDP-[^{14}C] glucose to glycogen[32]. A unit of activity is defined as the amount of enzyme that catalyzes the transfer of 1 µmol of glucose from UDP-glucose to glycogen per minute under the standard conditions of assay[33] (4.4 mM UDP-glucose and 6.7 mg/ml glycogen). The activity ratio of GS enzyme is defined as the ratio of activity measured in the absence of glucose-6-phosphate

to that measured in the presence of 7.2 mM glucose-6-phosphate . The activity ratio serves as a kinetic index of the activation state of the enzyme[34].

3. Catalytic mechanism of GS

GS is a metal ion independent retaining Leloir type (nucleotide sugar donor) glycosyl transferase[35,36]. Three different mechanisms have been proposed and are being investigated for the retaining class of glycosyl transferases, as follows.

i. S_N2 Mechanism

The S_N2 mechanism was originally proposed by Koshland[37], and involves two distinct nucleophilic displacement reactions and a covalent enzyme intermediate. It requires the presence of a catalytic nucleophile in the enzyme that initiates the reaction and a suitably positioned general acid/base that would assist the second nucleophilic attack (Figure 6). Enzymatic studies with human GYS1 led to the identification of the E-X_7-E motif that is conserved across the eukaryotic GS enzymes[38]. Mutation of the first conserved glutamate to alanine resulted in an inactive human GYS1 whereas mutation of the second glutamate to alanine resulted in 90% reduction of activity. Based on these results, the authors proposed that these conserved glutamate residues function as the nucleophile and general acid/base catalyst in GS enzymes.

ii. S_N1 Mechanism

The S_N1 mechanism for the retention of configuration at the anomeric carbon atom was first proposed by Philip for lysozyme[39]. The presence of an enzyme stabilized oxocarbenium ion intermediate could possibly shield and block the nucleophilic attack from the opposite face of the reaction center thus retaining the stereochemistry. However the free energy of the intermediate species and the associated transition states are very high. Since enzymes are believed to catalyze the reactions with lowest free energy intermediate species that would facilitate effective turnover, the S_N1 mechanism for retaining glycosyl transferase enzymes is not widely accepted by biochemists.

iii. S_Ni Mechanism

The mechanism of decomposition of alkyl chlorosulfites was the basis for the development of the S_Ni (internal return) mechanism. The leaving group undergoes decomposition leading to the production of a nucleophile that is held as an ion pair. The retention of stereochemistry is attributed to the high rate of decomposition of the intermediate ion pair and the nucleophilic attack by the product. A modified version of the internal return mechanism has been proposed for glycogen phosphorylase[40] and galactosyltransferase LgtC from *Nisseria Sp*[41]. The S_Ni mechanism for glycosyl transferases is partly based on the interaction between the departing phosphate of the sugar donor and the hydroxyl group of the acceptor[42,43].

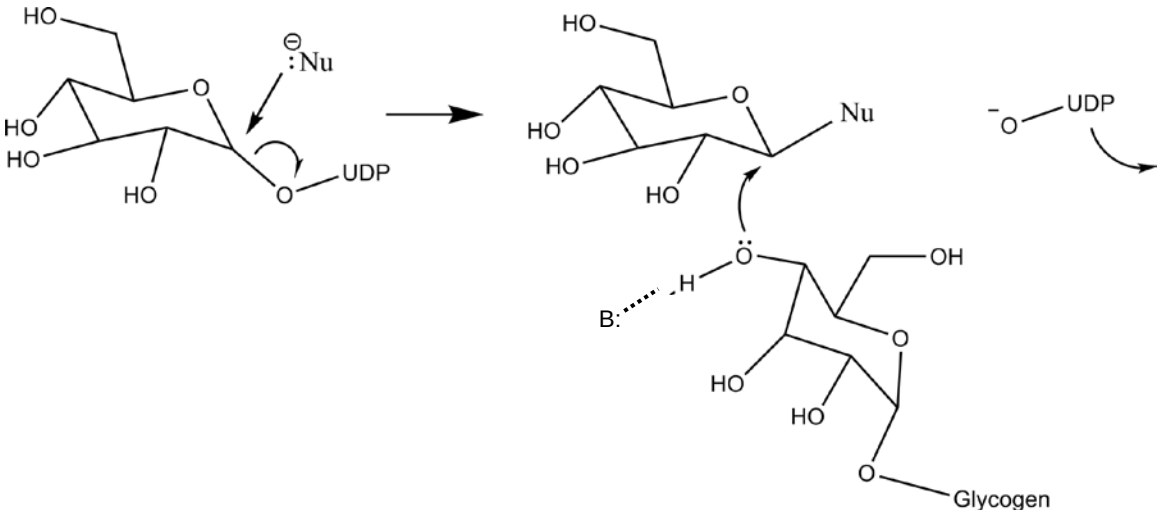

Figure 6. Proposed mechainsm of action for GS.
The S_N2 double displacement mechanism for retaining glycosyl transfer is depicted. The reaction involves two nucleophilic attacks, one by the catalytic nucleophile and the second by the activated 4'OH group of the acceptor. A suitably positioned general acid/base assists the second nucleophilic attack and the reaction involves a covalent intermediate. The S_N1 mechanism on the other hand is mediated by the presence of a charge separated oxocarbenium ion intermediate that blocks the second nucleophilic attack from the opposite face of the reaction center.

4. Influence of substrate on GS activity

Early studies of rabbit GYS1 with sugar acceptors of varying polymer length demonstrated that the $S_{0.5}$ of the sugar acceptor and the V_{max} of the enzyme changed significantly as the acceptor length was increased from three (maltotriose) to four (maltotetraose)[44]. When maltotetraose, maltohexaose and hydrolyzed amylose were used as acceptors there was no significant change in the enzymatic properties. However, when amylopectin subjected to varying levels of digestion by β-amylase was used as the substrate, a decrease in the V_{max} was observed as the chain length was decreased. When comparing maltose and β-amylase limit dextrin (two sugars in the outer chain) substrates, the V_{max} of the enzyme differed by four-fold, but the $S_{0.5}$ changed by five orders of magnitude. Based on these observations, the authors proposed the presence of distinct polysaccharide binding and catalytic sites in GS (Figure 7). Longer sugar polymers that occupy both the sites serve as better substrates for the enzyme. Further, Larner *et al.* hypothesized that the catalytic site is composed of a minimum of four identical sub-sites, the greater occupancy of which leads to greater enzymatic efficiency[44].

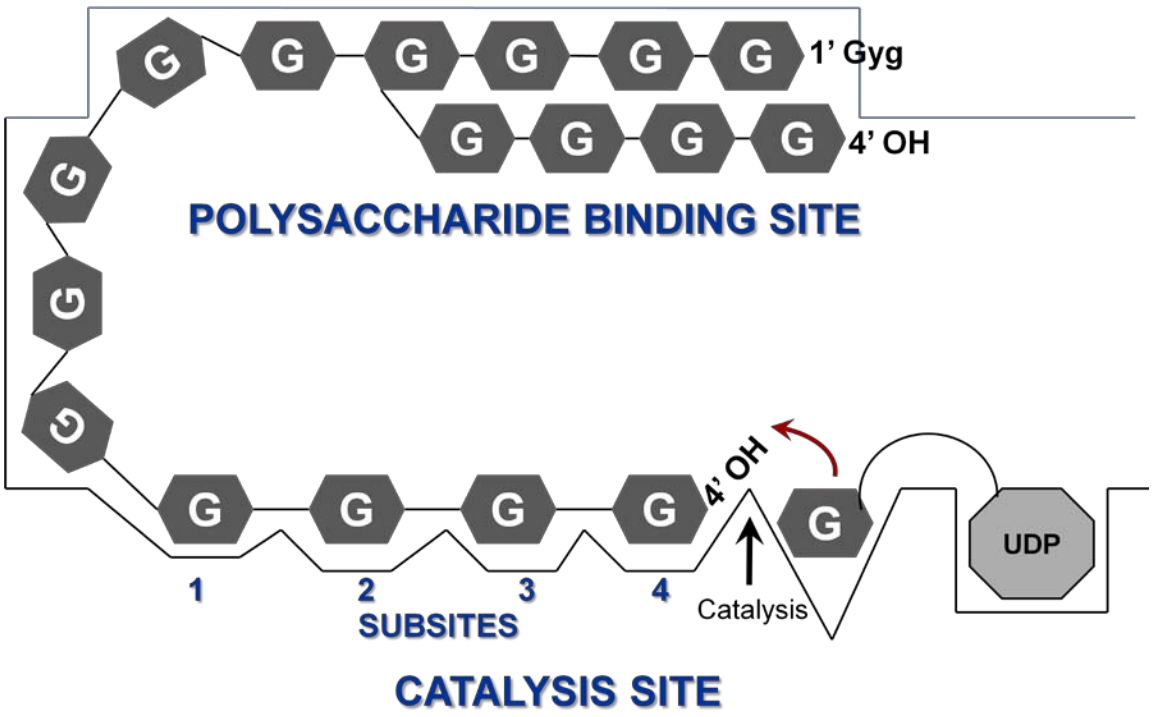

Figure 7. Model of active site and polysaccharide binding site of GS. Presence of separate catalysis and polysaccharide binding site in rabbit muscle GS. Sugar polymers that bind simultaneously to both the sites are better substrates for the enzyme. The active site is composed of four identical sub sites and acceptors that occupy a greater number of these sites act as better substrates.

5. Regulation of GS activity

The activity of eukaryotic enzymes is regulated by multiple mechanisms including covalent modification, allosteric activators and translocation within the cells. There are common regulatory themes – phosphorylation control and allosteric activation by glucose-6-phosphate, but the physiological responses that impinge on these regulatory controls often differ amongst different organisms and even between tissues of the same organism.

i. Regulation by covalent modification

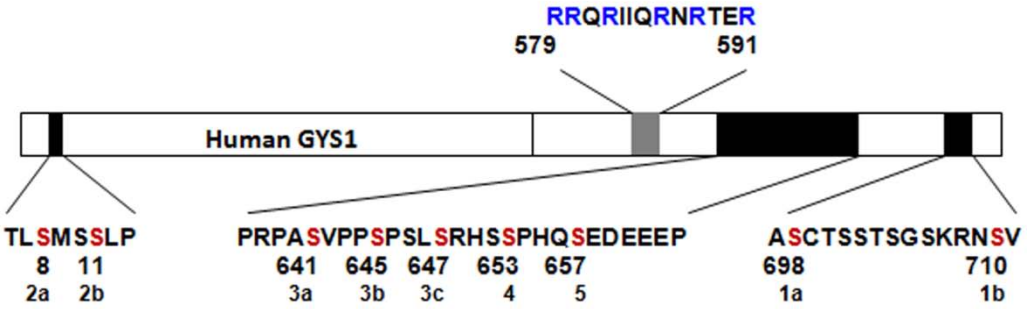

Figure 8. Schematic of human GYS1.
Schematic representation of human muscle GS. The phosphorylation sites and arginine cluster are represented as black and grey boxes the actual sequences are shown below. The phosphorylation sites and the conserved arginines are highlighted in red and blue respectively.

Hierarchal ATP-dependent phosphorylation by protein kinases of serine and threonine residues within conserved regions located at both the N- and C-terminal ends of the mammalian GS enzymes inhibits enzyme activity[45,46]. Phosphorylation of the rabbit GYS1 increases the $S_{0.5}$ value of the UDP-glucose substrate from 0.75 mM to 61 mM[47]. Mammalian GS enzymes have two potential phosphorylation sites in the N-terminal 20 amino acids and five to seven phosphorylation sites in the C-terminal 80-100 amino acids. Phosphorylation of the sites 3a-c has maximal effect on the enzymatic activity and studies with COS-M9 cells expressing the wild type and site 3a-c mutants have demonstrated that loss of these sites strongly correlates with increased glycogen accumulation[48,49]. Dephosphorylation of GS by type1 protein phosphatases (PP1) reverses the phosphorylation state of GS and a number of targeting subunits like R_{GL}[50]. G_L[51],

PTG[52] and PPP1R6[53] are involved in targeting PP1 to GS and glycogen. The importance of these targeting subunits for proper recruitment of PP1 is demonstrated by studies with knock-out mouse models. The skeletal muscle tissue of R_{GL} knock-out mice exhibit hyper phosphorylated GS which correlates well with the reduced glycogen levels in the muscle[54]. Decreased hepatic glycogen synthesis in insulin dependent diabetic rats has been linked to the decreased level of G_L espression in the liver of these animals[55].

Both N and C terminal phosphorylation sites of mammalian GS are involved in mediating insulin sensitivity[56] and impaired insulin regulation of GS in obesity and type 2 diabetes mellitus have been linked to the dysregulation of phosphorylation at sites 2, 2a and 3a-c of human muscle GS[57]. It has been reported that O-linked N-acetylglucosamine modification of GS restrains the enzyme in a glucose-6-phosphate dependent state and decreases the activation of the enzyme by insulin[58].

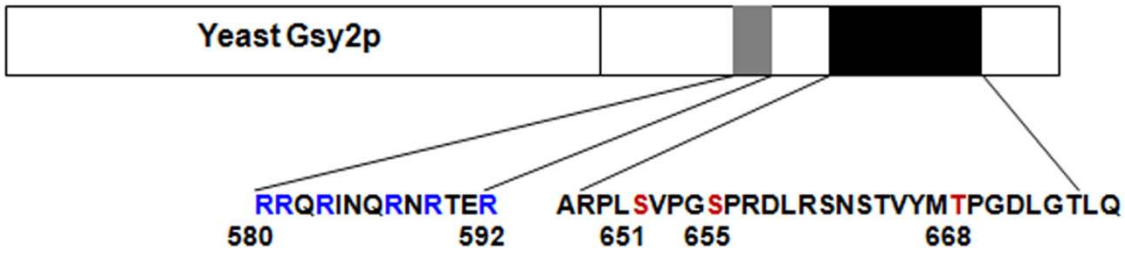

Figure 9. Schematic of Gsy2p.
The phosphorylation sites and arginine cluster are represented as black and grey boxes and the actual sequences are shown below. The phosphorylation sites and the conserved arginines are highlighted in red and blue respectively.

Yeast Gsy2p lacks the hierarchal phosphorylation control mechanisms and the N-terminal regulatory phosphorylation sites of the mammalian forms of GS but retains the inhibitory effects of C-terminal phosphorylation. In yeast, the C-terminal regulatory phosphorylation sites are residues Ser 651, Ser 655 and Thr 668[59]. Gsy2p truncated at residue 644 had a higher activity ratio and synthesized 4-fold more glycogen than its wild type counterpart[59]. Mutation of the Thr668 phosphorylation site to alanine resulted in a 35% increase in the activity ratio and an associated increase in glycogen accumulation[59]. In yeast, it has been established that phosphorylation of the GSy2p is mediated by the cyclin dependent protein kinase Pho85p[60] in association with the cyclin-like Pcl10p[61] subunit that activates and targets the kinase complex to the substrate. Studies with aspartate mutants of the phosphorylation sites in Gsy2p showed that the Thr 668 site was the most important site for activity control as only the mutant at this position showed decrease in activity[62]. However, the activity ratio of the T668D mutant did not show any considerable change suggesting that the aspartate

mutants are not actual phosphomimics for Gsy2p. Further evidence for the influence of Thr 668 on enzyme activity is the drastic decrease in the activity observed when the S651A/S655A double mutant was phosphorylated by the Pho85p/Pcl10p or Pcl8p complex (Roach and Wilson, unpublished). Similar to the mammalian enzymes dephosphorylation of Gsy2p involves a complex between the type-1 protein phosphatase Gac1p and its targeting subunit Glc7p[63,64]. Loss of GAC1 function leaves Gsy2p in the inactive from and decreases glycogen accumulation in the cells.

ii. Allosteric regulation

Inhibition of GS by phosphorylation can be overcome by the allosteric activator glucose-6-phosphate, which increases the V_{max} of Gsy2p by 2.5 fold[62]. In contrast, the effect of glucose-6-phosphate on the mammalian enzymes is primarily on substrate binding kinetics. It decreases the $S_{0.5}$ for UDP-glucose from >30 mM to ~50 µM with little effect on the V_{max}[47,65]. Alanine scanning mutagenesis of a conserved arginine rich sequence near the C-terminal portion of Gsy2p led to the discovery of its role in conferring sensitivity to glucose-6-phosphate and phosphorylation[62]. Triple mutation of the first three arginine residues to alanine (R580A/R581A/R583A) abrogated activation by glucose-6-phosphate and inhibition by phosphorylation. Mutation of the second set of three arginines also disrupted glucose-6-phosphate activation, but the enzyme could still be inhibited by phosphorylation. The same set of alanine mutations in the rabbit GYS1 enzyme also resulted in the loss of glucose-6-phosphate activation.

However, their sensitivity to inhibition by phosphorylation was swapped such that mutation of the N-terminal set of arginines retained inhibition by phosphorylation[66].

iii. Regulation by cellular translocation

Enzymes involved in glycogen synthesis change their intracellular localization in response to cellular glucose levels and this provides an additional mechanism of regulation. Insulin and glucose dependent redistribution of GS has been reported in skeletal muscle[67], adipocytes[68] and hepatocytes[69]. One study with GFP fused muscle GS expressed in C2L2 and COS-1 cells showed that the chimeric GS is localized near the nucleus at low glucose concentrations and upon increasing the concentration of glucose, it translocates to the cytosol and later adopts a punctate pattern of distribution[67]. A recent study with rabbit skeletal muscle has demonstrated that phosphorylation at sites 1b, 2 and 2a could regulate this redistribution[70]. Liver GS, on the other hand, is cytosolic even in the absence of glucose. However as the glucose levels are increased, an initial accumulation of GS at the periphery and subsequent distribution into the cytosol has been observed[69]. The redistribution of GS correlates well with the hepatic glycogen accumulation, which is initiated at the periphery and moves towards the interior as the glycogen deposits grow[14].

iv. Three state model for the regulation of yeast Gsy2p

Based on kinetic and mutational studies, a three state model has been proposed for the control of Gsy2p activity[62]. In the dephosphorylated state (I State), the enzyme exists in a intermediate activity state that has high sensitivity to glucose-6-phosphate (R State). Binding of glucose-6-phosphate converts this to the highest activity state. Phosphorylation of the intermediate form reduces the activity by 30 fold and decreases the sensitivity to glucose-6-phosphate by about 20 fold to generate the lowest activity state (T State). When exposed to saturating concentrations of glucose-6-phosphate, the phosphorylated form binds to glucose-6-phosphate and exhibits highest activity.

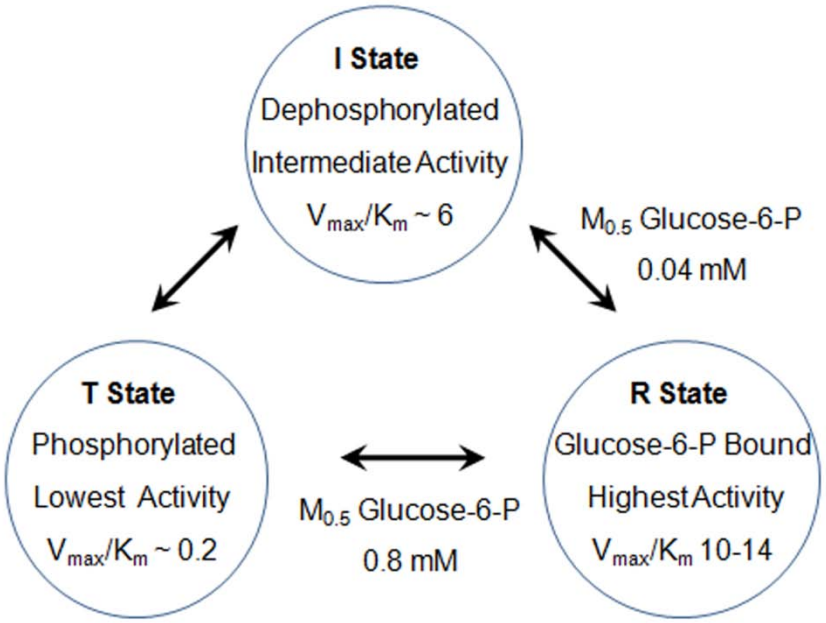

Figure 10. Three state model for the regulation of Gsy2p activity.
A three state model has been proposed for Gsy2p based on the kinetic and mutational studies. The dephosphorylated enzyme exhibits intermediate activity which upon binding to glucose-6-phosphate shows highest activity. Phosphorylation of the intermediate state decreases the activity by 30 fold. Figure adapted from *J Biol Chem.* 275, 27753-61, (2000).

6. Structural classification of GS

Sequence analysis, as well as more recent structure determinations, have grouped the GS enzymes within the GTB-fold of glycosyl transferases[71]. These structures are characterized by the presence of two Rossmann fold domains with a deep interdomain cleft in between that harbors the substrate-binding and catalytic sites. Within the larger group of GTB-fold enzymes, the GS enzymes are further subdivided into two families, GT3 and GT5. The bacterial and archaeal GS enzymes are grouped into the GT5 family and are not subjected to any known allosteric or covalent regulation. The eukaryotic enzymes are grouped into the GT3 family and are regulated through the allosteric activator glucose-6-phosphate and inhibitory phosphorylation. An additional point of distinction is that the bacterial enzymes use ADP-glucose as their sugar donor, whereas eukaryotic enzymes almost exclusively utilize UDP-glucose as their donor molecule. Archaeal enzymes are capable of using both ADP and UDP-glucose as sugar donors. To date, three dimensional structures have been determined for three members from the GT5 family - a monomeric *E.coli* enzyme[72,73], dimeric *Agrobacterium tumefacians* enzyme[74] and trimeric *Pyrococcus abyssi* enzyme[75]. However these structures have shed little light on the regulatory mechanisms controlling eukaryotic enzymes.

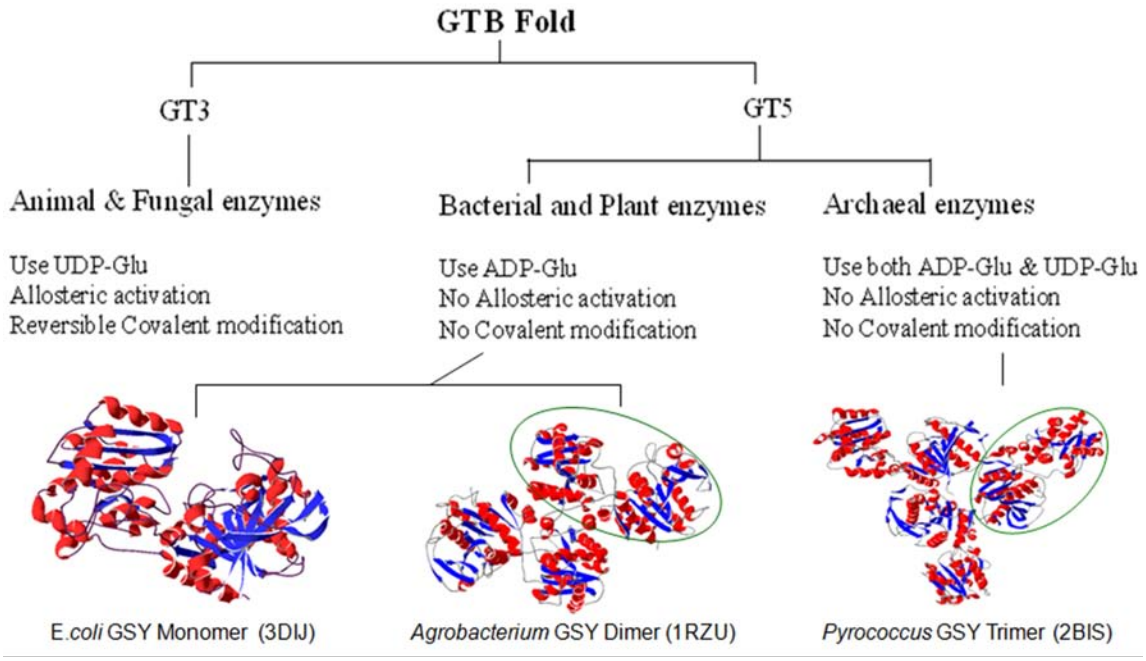

Figure 11. Schematic representation of the carbohydrate active enzymes (CAZy) classification of GS.
GS enzymes are grouped in the GTB fold and are further divided into two subgroups based on their regulatory properties. The structures of the three available GT5 enzymes are shown in ribbon representation and the individual monomers are highlighted in green.

C. Rationale and overview of the thesis research

The overall goal of this project is to understand the molecular mechanism of regulation of GT3 class of glycosyl transferase enzymes. We are studying the yeast Gsy2p as a model enzyme as it retains the regulatory mechanisms of the mammalian enzymes and shares 50% sequence identity to the mammalian enzymes.

The specific aims of this work include

a. Solving the three dimensional structure of Gsy2p in different activity states
b. Investigating the structural and functional consequence of the C-terminal phosphorylation of Gsy2p

Three major approaches have been implemented in the studies that include

a. Macromolecular X-ray crystallographic studies of the enzyme and enzyme complexes with activators and substrate analogs
b. Site-directed mutageneis and kinetic analysis of the mutants
c. Intein mediated peptide ligation approach to generate homogeneously phosphorylated forms of Gsy2p for kinetic and enzymatic studies

D. Theory of experimental methods used

1. Macromolecular x-ray crystallography

The purpose of this session is to provide a brief overview of the principles of X-ray crystallography and focus on (a) the multiple isomorphous replacement method of solving the phase problem and (b) the Patterson method[76] for solving the position of heavy atoms introduced into isomorphous derivatives.

i. Overview of the principles of macromolecular crystallography

Macromolecular structure determination by X-ray crystallography is based on the principle of diffraction of X-rays by the electrons associated with the protein molecules present in a crystal. Crystals consist of an ordered arrangement of protein molecules in a repeated arrangement within a unit cell, which is translated in three dimensions to form the lattice. Each unique crystal belongs to a particular space group, which specifies the symmetry relationship within the unit cell and is characterized by unique cell dimensions. Indexing the diffraction data refers to the process of determining these unit cell parameters and mathematically identifying each reflection by integral indices h, k and l. Each diffraction spot could be considered as reflection of the X-rays from an imaginary set of equidistant parallel planes identified by the indices h, k and l for the intercepts 1/h, 1/k and 1/l that the planes intersect with the unit cell edges a, b and c (Figure 12).

The reflections from the imaginary set of planes obey Bragg's law

$$2d \sin \theta = n \lambda$$

where d is the distance between the planes, θ is the angle between the plane and the incident rays, n is an integral whole number and λ is the wavelength of X-rays.

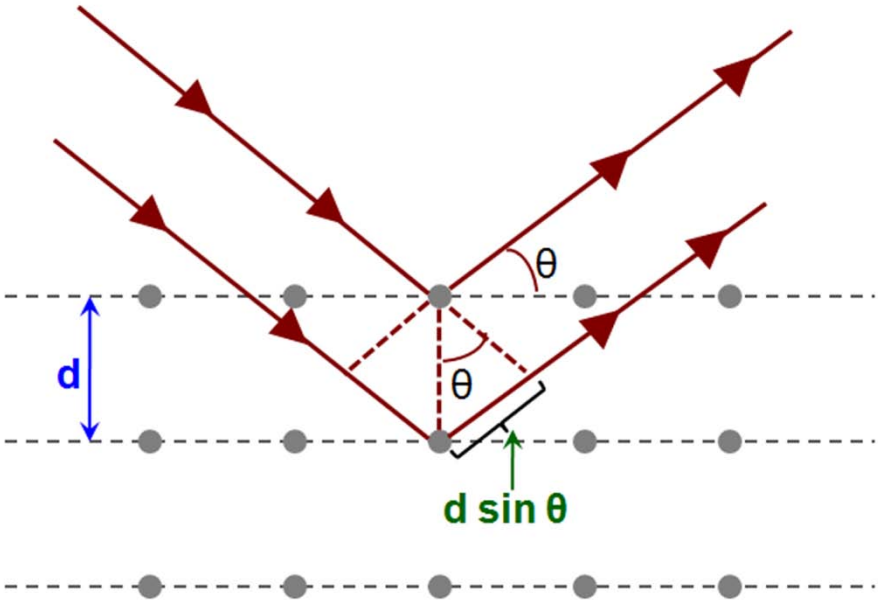

Figure 12. Bragg's law.
A two-dimensional crystal lattice and a set of imaginary planes is represented by the grid. X-rays incident on the crystal at an angle θ produce coherent diffraction (in phase) when the difference in the distance travelled (2d sinθ) is equal to an intergral number of the wavelength (nλ). Figure adapted from http://reference.iucr.org/dictionary.

The scattering of x-rays from the unit cells is dependent on the arrangement of the atoms in the unit cells. Each atom in the unit cell contributes to the diffraction by a structure factor that is characterized by a magnitude and phase angle, which depend on the atom type and scattering angle.

The structure factors can be described in the complex co-ordinate form as

$$A+iB = F (\cos \Phi + I \sin \Phi)$$

or is represented as a vector in the complex plane shown in Figure 13.

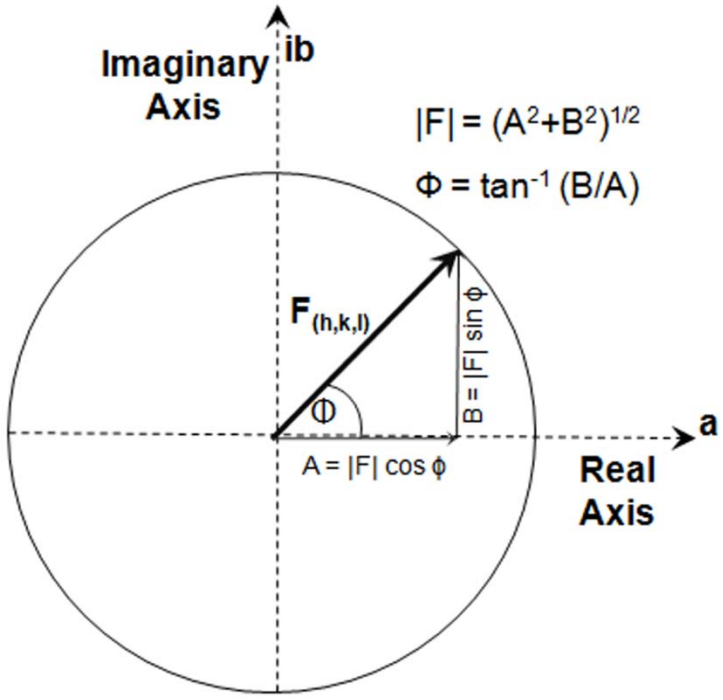

Figure 13. Vector representation of the structure factor.
The structure factor is represented as a vector A+iB / F (cos Φ+I sin Φ) in the complex plane. The amplitude or modulus of the structure facture is given by $(A^2+B^2)^{1/2}$ and the phase angle $\Phi = \tan^{-1} (B/A)$. (Figure adapted from X-ray structure determination – A practical guide. Stout and Jensen. Pg. 198)

The X-ray scattering from all the atoms of the unit cell are added to obtain the overall structure factor $F_{obs\ (h,k,l)}$, which is a vector describing the reflections from all the planes of index h,k and l. The overall structure factor is described as the summation of all the individual structure factors as

$$F_{obs\ (h,k,l)} = \Sigma\ f^{0}{}_{(i)} \cdot (\cos \varphi_{hkl} + i \sin \varphi_{hkl})$$

The magnitude of this structure factor is related to the intensity of the diffracted x-rays

$$I_{obs\ (h,k,l)} \propto | F_{obs\ (h,k,l)} |^2$$

The electron density of the molecules $\rho_{(x,y,z)}$ in the unit cell is calculated in the real space coordinates (x,y,z) using a Fourier synthesis of the structure factor amplitudes $F_{obs\ (h,k,l)}$, phase $\varphi_{(h,k,l)}$ and the unit cell volume (V) by the relation

$$\rho_{(x,y,z)} = V^{-1} \cdot \Sigma_{h,k,l} | F_{obs\ (h,k,l)} | \cos 2\pi [(hx+ky+lz) - \varphi_{(h,k,l)}]$$

Thus, to compute the electron density, we need both the amplitude and the phase information of the scattered X-rays. The intensity of the diffracted data (*i.e.* the square root of amplitude) is recorded when collecting data, but the phase information is completely lost. Hence it is impossible to determine the structure with the diffraction data alone. Several methods including multiple isomorphous replacement have been developed to solve the phase problem in crystallography.

ii. Phasing by multiple isomorphous replacement

Multiple isomorphous replacement, the most common method for experimental initio phasing was introduced by David Harker[77], but was first successfully applied in proteins to solve the structure of hemoglobin by Max Perutz[78]. The method involves the use of heavy atoms that are sufficiently electron dense as marker atoms to provide phase information. The heavy atom derivatives are prepared by soaking the native protein crystals with the heavy atom solution or by cocrystallizing the protein with the heavy atom. The presence

of additional electrons in the heavy atom will significantly change the scattered intensity of the X-rays. By comparing the native and the derivative data, we can compute the position of the heavy atom in the crystal and thus its contribution to the structure factor and the phase angle.

The structure factor of a isomorphous derivative (F_{PH}) is related to the sum of the structure factor of the protein (F_P) and that of the heavy atom (F_H) and is represented by a phase triangle.

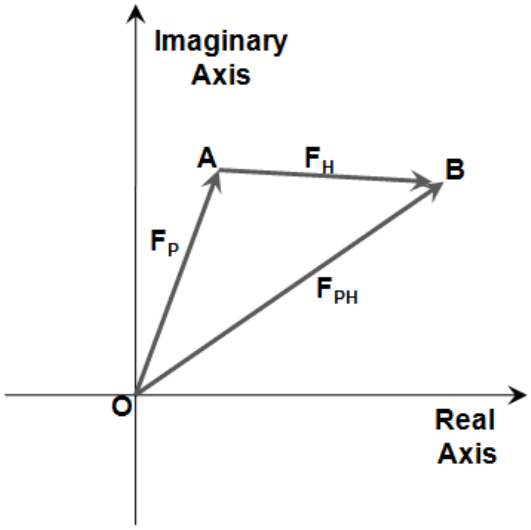

Figure 14. Phase triangle.
A vector diagram illustrating the contributions of the native protein (F_P) and the heavy atom (F_H) to the structure factor of the heavy atom derivative (F_{PH}). (Figure adapted from Protein Crystallography. Blundell and Johnson. Pg. 156)

The method of isomorphous replacement phasing is represented in the vector format by the Harker construction[76] (Figure 15). A circle (blue) with radius equal to the amplitude of protein $|F_P|$ centered at the origin (O) represents all the vectors that would be obtained with all the possible phase angles for F_P. If we now generate a new circle (green) with radius $|F_{PH1}|$ centered at a point defined by $-|F_H|$ from the origin, this would represent all the possible values for F_P that satisfy the equation $F_{PH1} = F_{H1} + F_P$ while agreeing with the measured amplitude $|F_{PH1}|$. The two points of intersection of these circles (H, L) represent the two possible values for the phase angles (F_P) that agree with both the measured amplitudes and with the heavy atom model. Thus, each derivative would give two possible phase values; one correct and one incorrect. This phase ambiguity is overcome by preparing a second different heavy atom derivative or using the anomolous scattering signal. The information from the second heavy atom derivative (red circle) in combination with the first will give the only phase choice that is consistent with all the information.

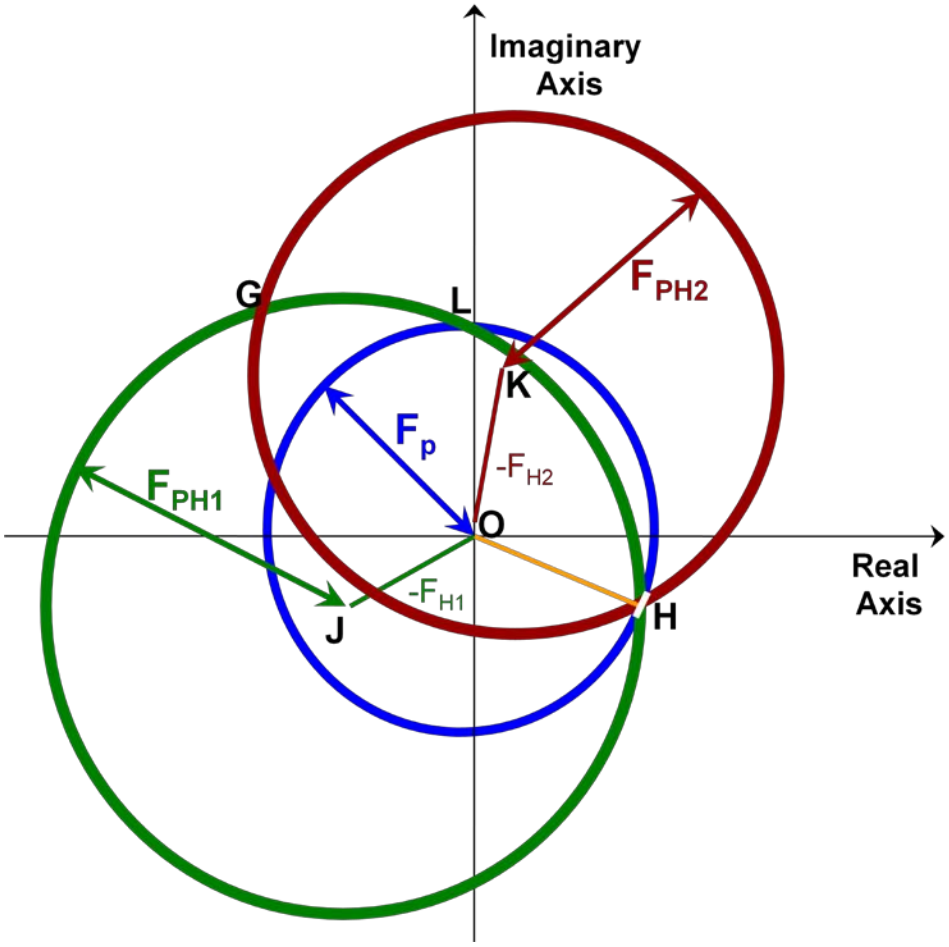

Figure 15. Harker construction for phase determination by the method of multiple isomorphous replacement.
Circles of radii F_P (blue), F_{PH1} (green) and F_{PH2} (red) are drawn with their centers at O, J and K respectively and the corresponding widths represent their individual errors in measurement. The vectors OJ and OK represent $-F_{H1}$ and $-F_{H2}$ the heavy atom contributions to the structure factor of the derivatives. The vector OH that defines the point of intersection of the three circles represents the magnitude and phase information of F_P. The error in phase estimation (lack of closure) is represented by the white bar and depicts the variation in possible phase assignment due to measurement or non-isomorphism error. (Figure adapted from Protein Crystallography. Blundell and Johnson. Pg. 161)

A perfect isomorphous derivative is one in which the crystal lattice is unperturbed by the addition of the heavy atom. In practice there is always some non-isomorphism between the native and derivative and this in combination with the errors is the measurement of the structure factor amplitudes results in the lack of closure of the phase triangle. The phase probability in the multiple isomorphous replacement method is the product of the individual normalized probabilities of each derivatives and is represented as

$$P(\varphi) = N \prod_j \exp\left(-\varepsilon_j^2 (\varphi) / 2E_j^2 \right)$$

where N is the normalization factor, $\varepsilon(\varphi)$ is the lack of closure of the triangle and E is the standard deviation. Thus MIR should yield one common solution with maximum probability among all the derivatives that represents the most probable phase and is calculated at the centroid of the probability distribution as

$$F_{best} = m |F_p| \exp(i \varphi_{best})$$

where m is the figure of merit and is the probability of φ_{best} being correct. If the error in phase angle at a given φ_i is defined as $\Delta\varphi_i = \varphi_{best} - \varphi_i$, m is the mean value of the cosine of error in the phase angle

$$m = \cos \Delta\varphi_i$$

A value of m = 1 corresponds to zero error in the phase angle measurement, while m = 0.5 corresponds to a phase error of 60°.

iii. Patterson function

The Patterson function was introduced by Arthur Lindo Patterson and is a modification of the electron density function where the structure factors originally represented by their amplitudes and phases are replaced by the square of their amplitudes that are proportional to the intensity of the diffracted X-rays.

$$P_{(u.v.w)} = V^{-1} \cdot \Sigma_{h,k,l} \mid F_{obs\ (h,k,l)} \mid^2 \cos 2\pi\ (hu+kv+lw)$$

where $P_{(u.v.w)}$ is the Patterson function and $F_{obs\ (h,k,l)}$ is the amplitude of the structure factor. The Patterson function is essentially a Fourier transform of the intensities rather than the structure factors. The Patterson function generates a centro-symmetric map of the inter-atomic vectors between atoms in the crystal structure. A peak at a position u,v,w in the Patterson map indicates that atoms exist in real space at positions x1,y1,z1 and x2,y2,z2 such that u = x1-x2, v =y1-y2, w = z1-z2. For a crystal with N atoms in the unit cell, the Patterson map will show N^2 peaks of which N will be the self vectors. The height of the peaks in the Patterson map is proportional to the product of the atomic numbers of the atoms involved and thus the Patterson map can be used for locating the heavy atoms in the molecule since the function is dominated by the square of the largest terms.

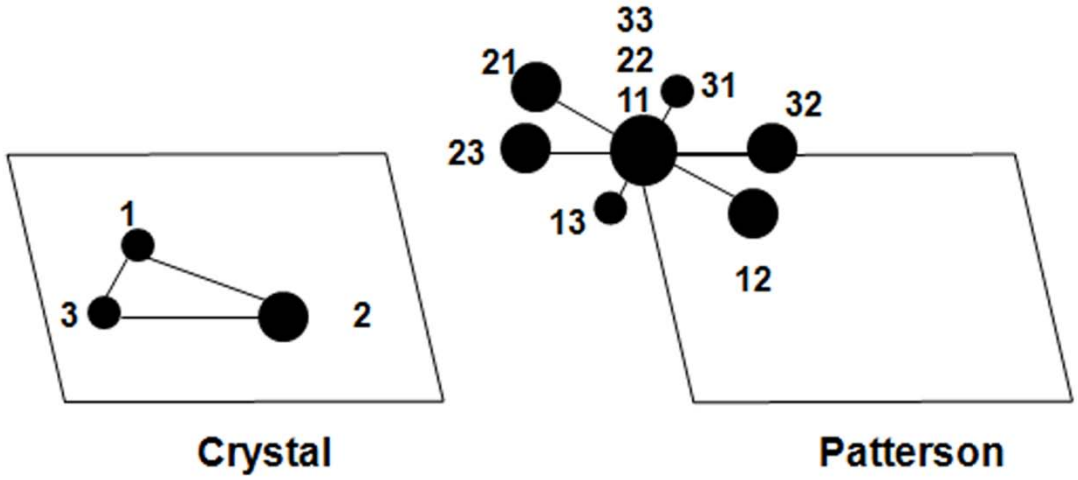

Figure 16. Patterson Function.
Representation of the Patterson function of a crystal with three atoms. The Patterson function is a map of all inter-atomic vectors between the atoms in the cell. (Figure adapted from Protein Crystallography. Blundell and Johnson. Pg. 138)

Harker showed that the interatomic vectors between equivalent atoms of the unit cell that are related by the symmetry operations produce peaks (Harker peaks) that are concentrated at certain planes or along certain lines determined by the known crystallographic symmetries of the unit cell[76]. These portions of the Patterson map are termed “Harker sections” and they provide an easy method for solving the atomic positions of electron dense atoms in the crystals through visual inspection of these sections.

iv. Solving the heavy atom position using the Patterson function

The primitive monoclinic $P2_1$ space group has one two-fold screw axis parallel to the *b* axis. The symmetry related general equivalent positions in the $P2_1$ space group are x, y, z and -x, ½+y, -z. The self-vectors in the Patterson for the $P2_1$ space group are calculated by subtracting one equivalent position from the other and are given by u, v, w = 2x, ½ , 2z. The Harker vector always occurs in the section defined by v = 1/2, which is generated by the 2-fold screw axis along b.

Thus for the $P2_1$ space group,

(2x, ½ , 2z)	=	(u, v, w)	or	(-u,- v,- w)
	or	(u+1, v, w)	or	(-u+1,- v,- w)
	or	(u, v, w +1)	or	(-u,- v,- w+1)
	or	(u+1, v, w +1)	or	(-u+1,- v,- w+1)

Using these relations, the equivalent positions of x and z in the real space co-ordinates can be calculated from the Patterson co-ordinates, but the y positions are unspecified. However since there is no unique origin point along the b axis in the $P2_1$ space group, the y value can be assigned arbitrarily. When two or more sites are present, the relative y value can be assigned using the cross-vectors between the sites[79]. The cross vectors are calculated by subtracting the positions of the individual sites and every pair of unique sites will have its own cross vector.

2. Intein mediated peptide ligation

Inteins are self-splicing protein elements and the process by which splicing occurs consists of four nucleophilic displacement reactions directed by residues in the inteins and the adjacent exteins[80]. The first step involves the N→S or N→O acyl shift and the N-extein is transferred to the first residue of the intein (Figure 17). This rearrangement is followed by a trans-esterification step and the N-extein is transferred to the side chain of a conserved cysteine or serine residue located at the intein C-extein junction. Succinimide formation involving a conserved asparagine residue within the intein breaks the amide bond between the intein and the extein and releases the intein. The final step in splicing involves a S→N acyl shift that facilitates the formation of a peptide bond between the N-extein and C-extein.

The process of protein splicing has been used in the expressed protein ligation or intein mediated peptide ligation approach that facilitates the ligation of synthetic peptides to recombinant proteins[80,81]. The recombinant protein is expressed in fusion with a chitin binding domain from *Bacillus circulans* and this enables the affinity purification of the fused protein with chitin beads. Similar to the protein splicing the first step in expressed protein ligation involves a N→S acyl shift of the recombinant protein to the cysteine residue of the intein. However the conserved asparagine is mutated to alanine and this prevents the release of the intein. Addition of a thiol reagent with a synthetic peptide that has

cysteine residue in the first position favors a trans thioesterification and S→N acyl shift, thus generating a semisyntheic protein.

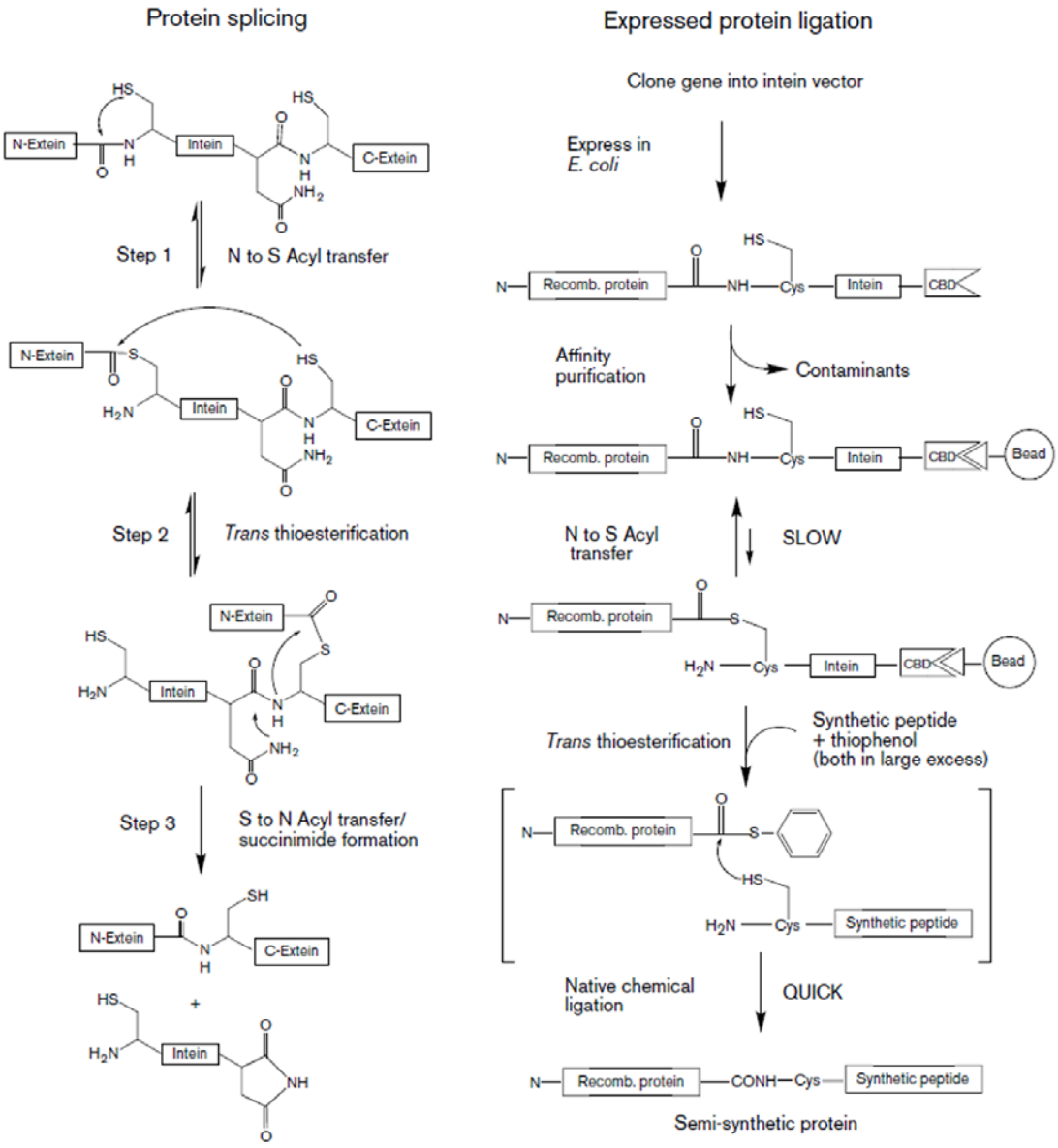

Figure 17. Mechanism of intein splicing.
Schematic representation of the nucleophilic reactions involved in the protein splicing and the expressed protein ligation method. Figure adapted from *Structure*. 6: 951-956 (1998)

METHODS

A. Gsy2p wild-type and mutant expression constructs

1. Site-directed mutagenesis

The wild-type, R580A/R581A/R583A and R587A/R589A/R592A Gsy2p His tagged construct in the pET28A *E.coli* expression plasmids were obtained from Roach lab. The other mutant constructs were created from these constructs using the PCR based site-directed mutagenesis approach[82]. Primers for amplification for site directed mutagenesis were designed and ordered from Integrated DNA Technologies. The PCR amplification of the plasmids was performed using the P*fu* Turbo polymerase from Stratagene. The parental plasmid in the amplified reaction mix was digested with Dpn I (New England Biolabs) and the Dpn I treated sample was introduced into competent *E.coli* DH5-α cells. Isolated transformed colonies were inoculated in LB broth containing kanamycin (50 µg / mL) and grown to OD_{600} of 0.7. The Promega WizardPlus SV miniprep kits were used for the purification of plasmid DNA from the transformed colonies. The mutations were confirmed by DNA sequencing of the plasmids at the DNA Sequencing Core Facility of IU School of Medicine.

2. Cloning of Gsy2p in the IMPACT vector

The yeast Gsy2p cDNA fragment encoding amino acids 1-640 was amplified from the original pET28A expression construct. The oligonucleotides for amplification were obtained from Integrated DNA technologies. The amplified cDNA was sub-cloned into the *Nde* I and *Sap* I sites of the IMPACT vectors pTXB1 and pTYB1 (New England Biolabs) and the ligated vectors were used to transform competent *E.coli* DH5-α cells. The transformed colonies were tested by the colony PCR method to check the insert size in the plasmid DNA. The colonies that showed the presence of the 2kb insert were inoculated in LB containing ampicillin (100 µg / mL) and were grown at 37°C to OD_{600} of 0.7. Plasmid DNA isolation from the cultures was performed using the Promega WizardPlus SV miniprep kits. In-frame ligation of the Gsy2p cDNA to the plasmids was confirmed by DNA sequencing at the DNA Sequencing Core Facility of IU School of Medicine using the T7 forward and MXe-intein reverse primers.

B. Peptide synthesis

Two 35-mer peptides corresponding to the C-terminal phosphorylation site of the Gsy2p were synthesized by the Peptide Synthesis Core of the Biochemistry and Biotechnology facility of Indiana University, School of Medicine. Subsequently two 49-mer peptides with additional 14 residues towards the C-terminal region were synthesized from Antagene.

The peptide sequences were as follows:

35mer Non-phosphopeptide

CKLKVARPLSVPGSPRDLRSNSTVYMTPGDLGTLQ

35mer Phosphopeptide

CKLKVARPLSVPGSPRDLRSNSTVYMT(PO_3)PGDLGTLQ

49mer Non-phosphopeptide

CKLKVARPLSVPGSPRDLRSNSTVYMTPGDLGTLQEVNNADDYFSLGVN

49mer Phosphopeptide

CKLKVARPLSVPGSPRDLRSNSTVYMT(PO_3)PGDLGTLQEVNNADDYFSLGVN

C. Expression and purification of Yeast Gsy2p

1. Protein preparation from pET28A constructs

The full-length yeast Gsy2p was expressed with the N-terminal His-tag **(**pET28A) in *E.coli* BL21 cells. Cultures were grown at 37°C to OD_{600} of 0.6 and induced with 0.1 mM IPTG at 16°C for 16 hours. Cells were collected by centrifugation and resuspended in lysis buffer containing 20 mM Tris pH 8.0, 300 mM NaCl, 10 mM imidazole, 2 mM Benzamidine, 1 mM β-mercapto ethanol (BME) and 0.1% TritonX-100. Cells were lysed by passage through French pressure at 17,000 psi and the lysate was clarified by ultracentrigugation at 35,000 rpm for 35 min. The clarified lysate was loaded on a Ni^{2+}-nitrilotriacetic acid-agarose (Qiagen) column pre-equilibrated with the lysis buffer. The column was washed to base line with the lysis buffer. Buffers with 25 mM and 50 mM

imidazole were used for additional wash steps to elute the non- specifically bound proteins. The bound Gsy2p was eluted with a linear gradient of 50 to 200 mM imidazole in the lysis buffer. The collected fractions were analyzed by running on a 10% SDS-PAGE gel and those containing Gsy2p were pooled and dialyzed against the Q-load buffer containing 50 mM Tris pH 8.0, 1 mM EDTA, 2 mM benzamidine and 1 mM BME. The dialyzed sample was loaded on to Q-sepharose fast flow (GE Healthcare) column pre-equilibrated with the Q-load buffer. The column was washed to base line with the Q-load buffer and bound protein was eluted with a linear gradient of 0-1M NaCl in the Q-load buffer. The collected fractions were analyzed by SDS-PAGE and fractions that contained >90% Gsy2p (visualized by Coomassie staining) were pooled and dialyzed against 20 mM Tris pH8.0 and 1 mM BME. The dialyzed protein was concentrated to 4-5 mg/ml at 4°C using either Amicon Centricon-30 or Pall Macrocon -50 centrifugal concentration devices. The concentrated protein was filtered through microfuge centrifugal filters, aliquoted to small volumes, flash frozen in liquid nitrogen and stored at -80°C.

2. Purification of Gsy2p core and semi-synthetic enzymes

The expression of the intein -chitin binding domain (CBD) fused Gsy2p from both the PTXB1 and PTYB1 were tested. Overnight induction at 16°C with 0.3 mM IPTG showed comparable levels of expression of the fusion protein from both the plasmids as analyzed by western blot using the anti-CBD antibody. Various thiol reagents including DTT, BME, MESNA and thiophenol at three

different temperatures - 4°C, 16°C and Room temperature (25°C) were tested for cleavage activity. The pTXB1 construct was chosen for the Gsy2p preparation as the protein purification from this construct resulted in reasonable protein yield (~5 mg from 2 L) in a reproducible manner. The detailed protocol for the expression and purification of Gsy2p from the PTXB1 consturct is given below.

E.coli ER2566 cells transformed with the pTXB1- yeast Gsy2p plasmids were grown in LB medium containing 100µg/mL ampicillin at 37°C to 0.6 OD_{600} and induced with 0.3 mM IPTG at 16°C for 16 hr. Cells were harvested and resuspended in lysis buffer (20 mM Tris pH 8.5, 500 mM NaCl, 2 mM EDTA, 2 mM Benzamidine, 1 mM PMSF and 0.1% Triton X-100). Cells were lysed by passage through the French press at 17,000 psi and the lysate was clarified by ultracentrifugation at 35,000 rpm for 35 min. The clarified lysate was loaded on to chitin resin (New England Biolabs) pre-equilibrated with buffer containing 20mM Tris-HCl and 500mM NaCl pH 8.5. The column was washed to base line with equilibration buffer (20 mM Tris pH 8.5, 500 mM NaCl, 2 mM EDTA, 2 mM benzamidine and 1 mM PMSF).

For the 1-640 truncated core enzyme, the column was treated with equilibration buffer containing 50 mM DTT and after standing at 25°C for 20hrs, the cleaved protein was eluted with the equilibration buffer. The semi-synthetic peptide ligated proteins were generated by treating the column with 2 ml of 2% v/v thiophenol in equilibration buffer followed immediately by 1 ml of 1 mM synthetic peptide and 2% thiophenol in equilibration buffer. After standing at 25°C for 20 hrs, the cleaved peptide ligated semi-synthetic protein was eluted with the

equilibration buffer. We observed that some amount of the cleaved Gsy2p was still bound to the chitin beads and hence an additional elution step with 50 mM maltose in the elution buffer was included. The eluted protein was subjected to dialysis against 20 mM Tris pH 8.0, 1 mM DTT and decreasing NaCl concentrations starting from 500mM to 0 mM, to remove the unligated peptide, salt and excess reducing agent. For the truncated protein all the dialysis were performed at 4°C. When the semi-synthetic proteins were generated, the first dialysis step was performed at 25°C to avoid precipitation of the thiophenol in the cold aqueous buffers. The purified enzymes were analyzed on a 10% SDS-PAGE gel and further concentrated using with Amicon Centricon-30 centrifugal concentration devices to a final concentration of 3 mg/mL. Ligation of the peptide was confirmed by running the protein with the His-tagged full length and 640 truncated control proteins on a 7% SDS-PAGE gel. The concentrated protein was aliquoted to small volumes, flash frozen in liquid nitrogen and stored at -80°C.

D. Crystallization of Gsy2p

1. R580A/R581A/R583A crystals

Crystals of Gsy2p – R580A/R581A/R583A mutant were grown in hanging drops by combining 2µl protein at 3mg/ml with 2µl reservoir solution containing 100 mM Tris-HCl pH 8.0-8.5, 200 mM Li_2SO_4 and 18-22% PEG 3400. Physically twinned six-sided plate clusters of crystals were obtained in 3-5 days. These

crystals were used for making seed stock solutions using the seed bead kit from Hampton Research. The seed stock solutions were prepared in the crystal stabilizing solution containing 100 mM Tris - HCl pH 8.3, 200 mM Li_2SO_4 and 25% PEG 3400 and serial dilutions of the seed stocks were made with the same stabilizing solution. Both the streak seeding and microseeding techniques were used and the microseeding gave reproducible results. Trapezoidal crystals of Gsy2p appear in 5 days and grow to a length of 200 µm in two weeks.

Glycerol to a final concentration of 20% was used as the cryoprotectant for freezing the crystals. The crystals were incubated in the mother liquor with 6% glycerol for 16 – 20 hours before harvesting. The treated crystals were transferred quickly to the mother liquor with 10% and 20% glycerol and were flash frozen at -180°C in the gaseous nitrogen stream on the X-ray machine.

2. R589A/R592A glucose-6-phosphate Co-crystals

Screening for crystallization conditions of the GSy2p – R589A/R592A mutant, with and without glucose-6-phosphate (25 mM) was performed using the commercially available crystal screens from Hampton Research and Emerald Biosystems. The hanging drop vapor diffusion method was used for the screening and the drops were set by combining 2 µl protein at 3 mg/mL with 2 µl of the crystallization solution. Crystal plates were incubated for 1 week at room temperature. One condition (100 mM Bis-Tris pH 6.5, 25% PEG 300) with which the R589A/R592A double mutant crystallizes only in the presence of glucose-6-phosphate was discovered and was chosen for further optimization.

Co-crystals of the yeast Gsy2p – R589A/R592A mutant with glucose-6-phosphate were grown in hanging drops by combining 2 µl protein at 3 mg/mL and 50 mM glucose-6-phosphate with 2 µl reservoir solution containing 100 mM Bis-Tris pH 6.2-6.5, and 18-22% PEG 300. Cryoprotection worked best if the crystals were first dehydrated by replacing the mother liquor with a solution containing the same buffer containing 30% PEG 300 and re-incubating the plates for 48-60 hours before harvesting. The stabilized crystals were quickly coated with immersion oil before freezing at -180°C in the gaseous nitrogen stream.

3. Heavy atom and ligand soaks

The trapezoidal R580A/R581A/R583A crystals were transferred to the crystallization solution with the same buffer and 200 mM Li_2SO_4 but 30% PEG 3400 (stabilizing solution) and incubated for 12 hours. Heavy atom derivatives of the stabilized R580A/R581A/R583A crystals were prepared by soaking the crystals in the crystal stabilizing solution containing 1 mM Ta_6Br_{12} [83] for 6-12 hours. The UDP and maltooctaose complexes with Gsy2p were prepared by soaking the native crystals in the crystal stabilizing solution containing 20 and 50 mM of the respective compounds. All the ligand and heavy metal soaked crystals were cryoprotected and frozen in the same way as the native R580A/R581A/R583A crystals.

E. Data collection, processing, structure solution and refinement

1. Data collection

i. R580A/R581A/R583A data sets

The 2.9 Å native data set and the tantalum derivative data sets were collected at the 19-ID beamline operated by the Structural Biology Center at the Advanced Photon Source located at the Argonne National Laboratory. Fluorescence scans were performed on the tantalum derivatives and data collection was initiated at the peak wavelength for these derivatives; 1.25 Å. The 3.0 Å native data set, the UDP, as well as the maltooctaose complex data sets were collected on beam line 23-ID at the Advanced Photon Source operated by the General Medicine and Cancer Institute Collaborative Access Team. The data collected at the 19-ID beamline were indexed, integrated and scaled using the program package HKL3000 and the data collected at the 23-ID beamline were processed with HKL2000[84].

ii. R589A/R592A glucose-6-phosphate co-crystals data set

The glucose-6-phosphate co-crystal dataset was collected on beam line 23-ID at the Advanced Photon Source operated by the General Medicine and Cancer Institute Collaborative Access Team and the data was processed with HKL2000[84].

2. Structure solution, model building and refinement

i. R580A/R581A/R583A native 1 structure

The R580A/R581A/R583A mutant crystallizes in the primitive monoclinic $P2_1$ space group with two similar but distinct cell dimensions. The structure of the native1 R580A/R581A/R583A was solved using the method of multiple isomorphous replacement from two distinct tantalum bromide cluster (Ta_6Br_{12}) derivatives, combined with 4-fold averaging, solvent density modification and iterative partial model phase combination to extend the resolution of the electron density maps from 5.5 Å to 3 Å. The intial refinement of the structure was performed with the program Refmac5 and subsequent refinements were done with PHENIX, which improved the R values by 3-5%. The conjugate gradient matrix refinement protocol utilized the maximum likelihood target function and NCS restraints for each subunit along with tightly restrained individual temperature factor values.

ii. R580A/R581A/R583A native 2 and ligand complex structures

The native 2 and ligand complex structures were solved by molecular replacement method using the program AMORE and the refined native 1 structure as the search model. Rigid body refinement of the AMORE solutions was performed using the program Refmac5 defining each monomer as the rigid body. This was followed by restrained refinement with NCS restraints for individual subunits. For the ligand complexes, the resulting maps were used to

manually dock the respective ligands using the program Coot and the merged structures were subjected to further conjugate gradient matrix refinement with Refmac5 and PHENIX using the maximum likelihood target function. The monomer libraries for the UDP and maltotetraose molecules were generated from the glycogenin (PDB id: 1LL2) and pullulanase (PDB id: 3FAX) structures respectively. For the 3.5 Å UDP-complex structure, a single overall B-factor for the protein and a separate overall B-factor for the ligands was refined.

iii. R589A/R592A glucose-6-phosphate complex structure

The R589A/R592A and R580A/R581A mutants cocrystallized with glucose-6-phosphate in the I222 space group. The structure of R589A/R592A mutant was solved my molecular replacement using the program AMORE and the R580A/R581A/R583A monomer as the search model. Each monomer was defined as a rigid body for the rigid body refinement in Refmac5. Subsequent restrained refinements were performed with Refmac5 and PHENIX and all refinement protocols utilized NCS restraints for each domain (N-terminal domain, C-terminal domain and long helical domain) in each subunit.

F. Structure analysis

1. Protein surface analysis

The accessible surface area of the protein was calculated with the Areaimol[85] program of the CCP4 suite. The analysis of the protein interfaces were performed with the Protein Interfaces, Surfaces and Assemblies (PISA)[86]

service at the European Bioinformatics Institute. The buried surface area of the tetrameric interface was calculated by treating each dimer of the dimer-of-dimers assembly as a single polypeptide.

2. Domain rotation analysis

The LSQ-superpose program of the CCP4 suite was used for analyzing the domain rotations in the oligomers. For the inter-domain rotation measurement in R580A/R581A/R583A, the C-terminal domains of subunit B and D were aligned and the rotational matrices for the N domain were generated using the above aligned structures.

The overall conformational change induced by glucose-6-phosphate was determined by aligning the long helical domains of the AC dimer of the R580A/R581A/R583A and R589A/R592A structures. The changes in the inter domain rotations were calculated by superposing the long helices of the monomer s and calculating the rotational matrices for the N and C domain alignments.

G. GS activity measurement and data analysis

1. Preparation of treated glycogen for GS assay

The mixed bed resin TMD-8 hydrogen and hydroxide form (Sigma-Aldrich) was washed with 5 column volumes of distilled water. A freshly prepared 10% solution of rabbit liver glycogen type III (Sigma Aldrich) was applied to the column and washed with distilled water to elute glycogen. Glycogen containing fractions

(turbid appearance) were pooled and mixed with ice-cold 95% ethanol to a final concentration of 66% ethanol. Two drops of 1 M NaCl solution was added to facilitate glycogen precipitation and the precipitated glycogen was collected by centrifugation at 8,000 rpm for 20 min at 4°C. The glycogen precipitate was further washed with 95% ethanol and recollected by using a similar centrifugation method. The precipitated glycogen was set to air dry and the dried glycogen was ground to coarse powder before dissolving it to 8% w/v solution with distilled water.

2. GS assays

i. Standard assay

Glycogen synthase activity was determined by the method described by Thomas *et al.*[32] and quantified my measuring the amount of ^{14}C-glucose incorporated from UDP-[^{14}C]-glucose into glycogen and the assay reactions were at 30°C for 15 min. The 10X dilution buffer contains 500 mM Tris-HCl, 200 mM EDTA, 250 mM KF pH 7.8. 1X GS dilution buffer was prepared by a 10-fold dilution of the 10X buffer and inclusion of treated glycogen and DTT to a final concentration of 0.1% and 1mM respectively. For the activity determination, the Gsy2p stock solution was diluted to 0.02 mg/ml with cold dilution buffer containing 20 mM Tris pH 8.0 and 1 mM DTT. The protein was further subjected to 1:5 dilution with 1X GS buffer and incubated at RT for 10 min. 25 µL of the diluted protein sample was added to 50 µL reaction mix (prewarmed for 10 min at

30°C) and incubated at 30°C for 15 min following which 62.5 µL was spotted on the filter paper. After extensive washing with 66% ethanol and drying the ^{14}C counts were measured with a scintillation counter. GS activity was measured under the standard assay conditions with 6.7 mg/ml glycogen and 4.4 mM UDP-glucose in the absence or presence of 7.2 mM of glucose-6-phosphate.

ii. Sulfate titration

The sulfate titration was done under the standard assay conditions with six different concentrations of Li_2SO_4 between 0-1 M. A stock solution of 2 M Li_2SO_4 pH 7.8 was prepared in 10 mM Tris-HCl pH 7.8 and this was used for preparing the reaction mixes for the sulfate titration of Gsy2p.

iii. Glucose-6-phosphate titration

For the glucose-6-phosphate activation, the concentration of glycogen and UDP-glucose in the reaction mixes were 6.7 mg/ml and 5 mM respectively and the activity was measured in the absence and presence of six concentrations of glucose-6-phosphate in the range of 0.03-3 mM.

iv. UDP-glucose titration

The UDP-glucose kinetics was measured with reaction mixes containing 6.7mg/ml of glycogen, and six concentrations of UDP-Glucose in the range of 0.075-5 mM, either in the presence or absence of saturating concentrations of glucose-6-phosphate.

v. Glycogen titration

For the glycogen titration the enzymes were diluted in 1X GS dilution buffer that had no added glycogen. The assays were done with 4.4mM UDP-glucose in the reaction mixes either in the absence or presence of saturating concentrations of glucose-6-phosphate. Nine different concentrations of glycogen in the range of 0.03-12mg/ml were used.

3. Kinetic data analysis

Enzyme kinetic data were analyzed using the SigmaPlot software package. The UDP-glucose kinetics data were fitted to the Michaelis – Menten equation $V=V_{max}*[S]/(Km+[S])$. The glucose-6-phosphate activation data were fitted to a three parameter rectangular hyperbola $V=V_0+ a*x / (b+x)$, where V_0 is the rate in the absence of glucose-6-phosphate, x is the concentration of glucose-6-phosphate, a is the maximal velocity and b is the concentration of glucose-6-phosphate required for half maximal activation ($M_{0.5}$). The kinetic constants were determined from these fits. The glycogen titration data was fitted to the standard Michaelis-Menten equation $V= a*x / (b+x)$, where a is the V_{max} (glycogen varied) and b is the concentration of glycogen required for half maximal activity ($S_{0.5}$).

4. Dephosphorylation of phosphopeptide ligated semi-synthetic Gsy2p by protein phosphatase treatment

The phosphopeptide ligated semi-synthetic protein stock solution (3mg/ml) was diluted 100-fold with 20 mM Tris pH 7.0. The diluted protein was treated with protein phosphatase PP1γ (final concentration 50 µg/ml) for 1 hr at 30°C in the presence of 0.2 mM $MnCl_2$. The reaction was terminated by a 10-fold dilution of the reaction mix with the 1X GS buffer wherein the EDTA present in the 1X buffer chelates the Mn ions required for the catalytic activity of the phosphatase enzyme.

RESULTS

A. Expression and purification of recombinant Gsy2p

1. Purification of His-tagged full length Gsy2p

His-tagged Gsy2p was expressed in *E.coli* BL21-DE3 cells and purified from the cell lysate by Ni^{2+} affinity chromatography followed by ion exchange chromatography on Q-sepharose. Typical protein yields from 6L bacterial culture were 25-30 mg and the yields were consistent amongst the wild type and different mutants. The final Q-pool was greater than 90% pure as analyzed by Coomassie staining on a SDS-PAGE gel (Figure 18). All protein concentrations were measured using the Biorad protein assay kit with bovine serum albumin as the standard.

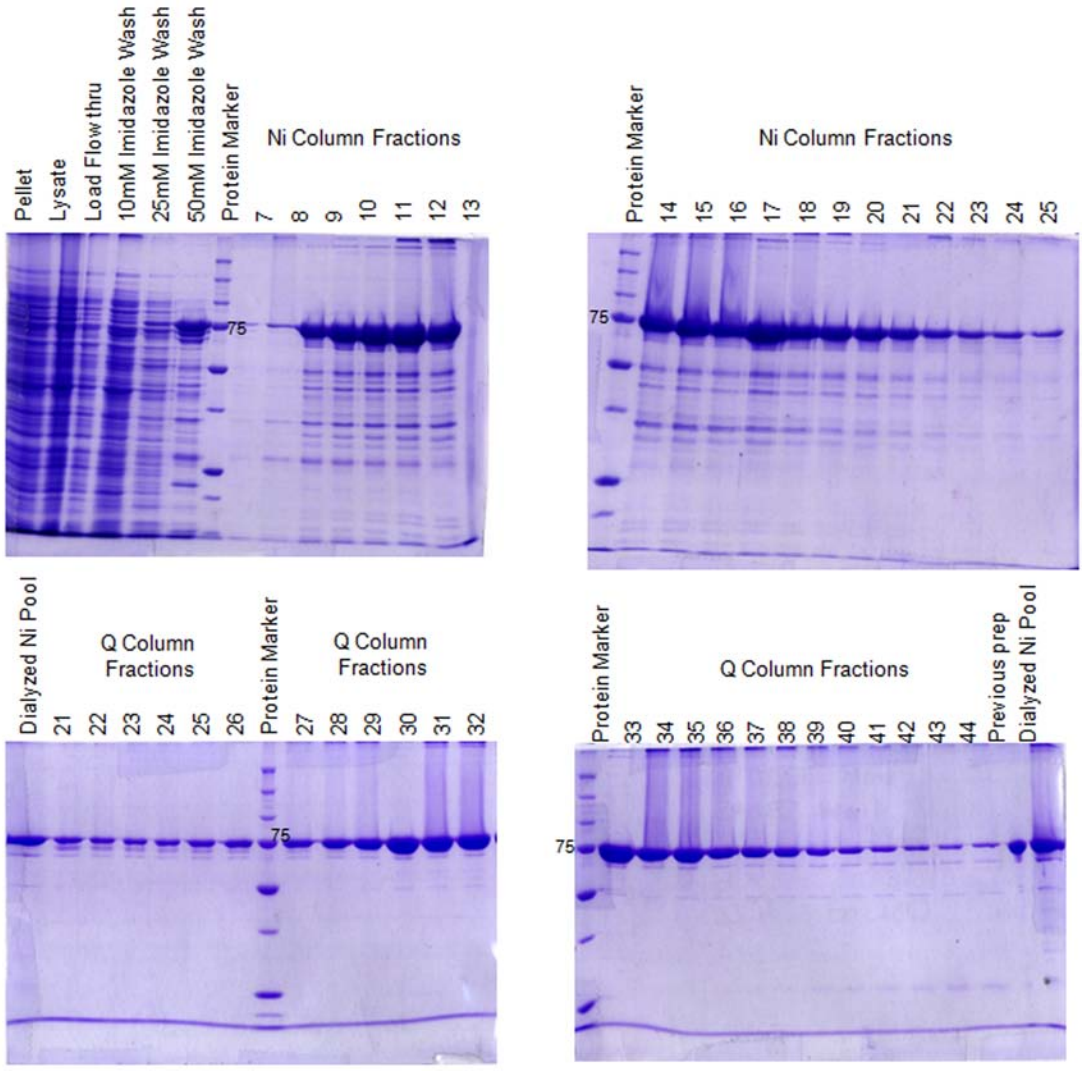

Figure 18. Representative SDS-PAGE gels of His-tagged Gsy2p Prep.
Top panel. Ni-NTA Column purification of R580A/R581A/R583A Gsy2p. Bottom Panel. Q-Sepharose column purification of of R580A/R581A/R583A Gsy2p. Typical yield from 6L prep was ~25-30 mg protein.

2. Purification of truncated and semi-synthetic Gsy2p

The Gsy2p Mxe GyrA Intein-CBD fusion protein was expressed in *E.coli* ER2566 cells and affinity purified from cell lysate using chitin resin. Overnight cleavage with 50mM DTT at room temperature resulted in ~ 6 mg of truncated protein from 2 L culture. Cleavage and coupling with 2% thiophenol and 1 mM synthetic peptide gave a typical yield of ~3-5 mg of semi-synthetic protein from 2L culture. Analysis of the His-tagged full length, 640 truncated and peptide ligated semi-synthetic proteins on a 7% SDS-PAGE gel showed clear shift in the position of the protein bands (Figure 19).

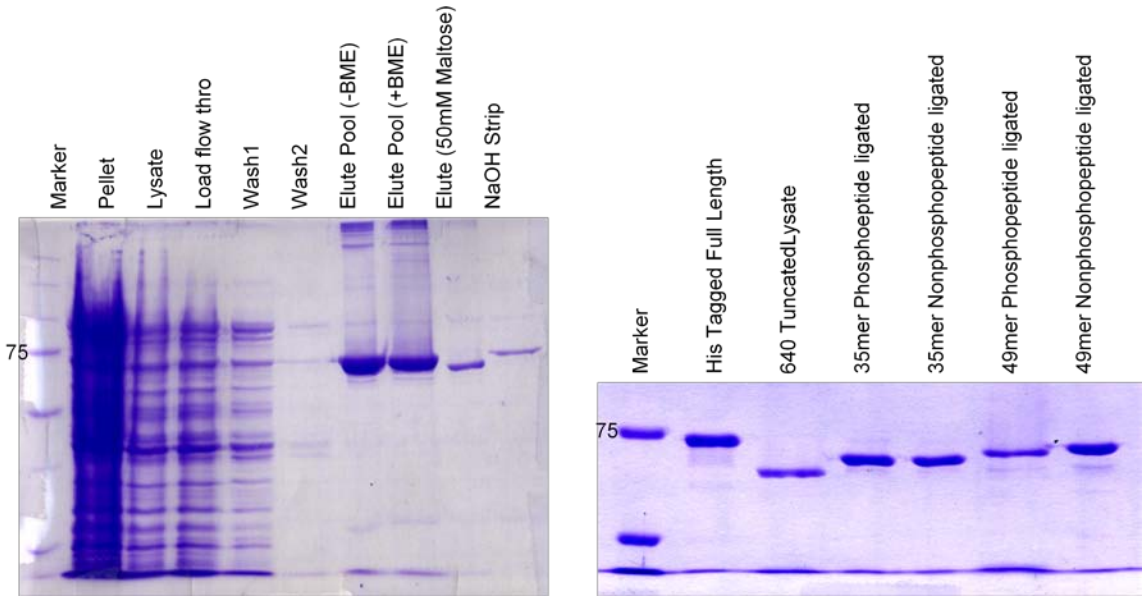

Figure 19. Semi-synthetic Gsy2p Prep and Mobility shift of Gsy2p.
Left Panel. 8% SDS-PAGE gel of Chitin Resin Purification and Coupling of 35mer Phosphopeptide to wild type Gsy2p. Right Panel. 7% SDS-PAGE gel of Wild type, 640 Truncated and Synthetic peptide ligated Gsy2p.

B. Specific activity, activity ratio and UDP-glucose and glucose-6-phosphate kinetics of Gsy2p

All the kinetic measurements are the average from at least three separate experiments where assays were performed in duplicate for each individual experiment. With the exception of the semi-synthetic single (R580A) and double (R580A/R581A) arginine mutants, the results represent assays from at least two different protein preparations. The wild type recombinant Gsy2p has a basal activity of 1.03 $mmol.min^{-1}$ with an activity ratio of 0.6. Consistent with the earlier work, both the triple alanine mutants are insensitive to glucose-6-phosphate activation and exhibit 35 to 45% decrease in the basal activity (Table 1).

To delineate the role of the conserved arginines in mediating inhibition by phosphorylation, we used the expressed protein ligation approach to ligate synthetic phosphopeptides and non-phosphopeptides to constructs of Gsy2p truncated at residue 641 with either wildtype sequences or mutated sequences within the arginine cluster. The sequence for the fusion peptides was identical to the wild type Gsy2p for residues 643 through 676 (35mer) or 643 through 690 (49 mer) with or without a phosphothreonine at position 668. Since an N-terminal cysteine residue is required for ligating synthetic peptides in the intein mediated peptide ligation approach, a non-conserved lysine in the wild type enzyme at position 642 was replaced by cysteine in the semi-synthetic proteins. Gsy2p truncated at residue 641 and the control non-phosphorylated peptide-ligated enzymes exhibit activities and activity ratios similar to the recombinant wild type

enzyme. In contrast, introduction of a single phosphothreonine at position 668 in the peptide-ligated construct inhibits the enzyme in a manner comparable to the wild type enzyme phosphorylated by the Pho85p/Pcl10p complex[62]. Ligation of the 35mer phosphopeptide to the wild-type enzyme decreases the basal activity by 70% and reduces the activity ratio about four-fold. Ligation of the 49-mer phosphopeptide decreases the basal activity of the enzyme by 90%. Incubation of the phosphopeptide ligated samples with the protein phosphatase PP1 restored the activity ratio of the enzymes to that of the control non-phosphopeptide ligated enzymes.

The influence of phosphate on substrate and activator binding was studied by titrating Gsy2p with UDP-glucose and glucose-6-phosphate (Table 2). In both the basal and activated states, the wild-type full-length and truncated forms exhibit similar K_M [UDP-Glu] and $M_{0.5}$ values for glucose-6-phosphate. However the K_M [UDP-Glu] of the phosphorylated form in the basal state is five times higher than that of the wild type enzyme and upon activation by glucose-6-phosphate, is restored to the wild-type values. The concentration for half maximal activation ($M_{0.5}$) by glucose-6-phosphate for the phosphorylated enzyme was three times that of the wild type enzyme.

Table 1. Specific activity and activity ratio of Gsy2p

Enzyme	Synthase Specific activity (mmol. min^{-1})		Activity Ratio -G6P/+G6P
	No G6P	7.2 mM G6P	
Wild Type	1.03 ± 0.03	1.73 ± 0.03	0.60 ± 0.01
Wild type-640 Truncated	1.25 ± 0.04	1.95 ± 0.05	0.64 ± 0.01
Wild type-35mer Nonphosphopeptide ligated	1.12 ± 0.02	1.77 ± 0.05	0.63 ± 0.01
Wild type-35mer Phosphopeptide ligated	0.27 ± 0.02	1.96 ± 0.04	0.14 ± 0.01
PP1 Treated Wild type-35mer Phosphopeptide ligated	1.18 ± 0.02	2.04 ± 0.08	0.58 ± 0.01
Wild type-49mer Nonphosphopeptide ligated	1.08 ± 0.02	1.78 ± 0.02	0.61 ± 0.02
Wild type-49mer Phosphopeptide ligated	0.10 ± 0.01	1.92 ± 0.04	0.05 ± 0.01
PP1 Treated Wild type-49mer Phosphopeptide ligated	1.34 ± 0.03	2.17 ± 0.02	0.62 ± 0.01
R580A/R581A/R583A	0.55 ± 0.02	0.51 ± 0.03	1.09 ± 0.02
R587A/R589A/R592A	0.66 ± 0.01	0.58 ± 0.02	1.14 ± 0.05

GS activity was measured in the presence of 6.7mg/ml of glycogen and 4.4 mM UDP-glucose in the absence and presence of 7.2 mM glucose-6-phosphate. The errors indicate the S.E.M. from three independent experiments

Table 2. UDP-glucose and glucose-6-phosphate kinetic parameters of Gsy2p

Enzyme	K_M [UDP-Glucose] mM		$M_{(0.5)}$ G6P mM
	-G6P	+G6P	
Wild type-Full length	0.54 ± 0.03	0.28 ± 0.02	0.04 ± 0.004
Wild type-641 truncated	0.62 ± 0.05	0.35 ± 0.03	0.05 ± 0.006
Wild type-T667 35mer Phosphopeptide ligated	2.4 ± 0.40	0.45 ± 0.06	0.13 ± 0.004
Wild type-T667 49mer Phosphopeptide ligated	ND	ND	0.12 ± 0.002

ND – Not Determined

UDP-glucose titration. GS activity of the wild-type and Thr668 phosphopeptide ligated semisynthetic wild type Gsy2p were measured in the presence of 6.7 mg/ml of glycogen and six concentrations of UDP-glucose in the range of 0.075-5mM in the absence or presence of 7.2 mM glucose-6-phosphate. The data was fitted to the Michaelis – Menten equation $V=V_{max}*[S]/(K_m+[S])$ to extract the K_M values.

Glucose-6-phosphate titration. GS activity was measured in the presence of 6.7 mg/ml of glycogen, 5 mM UDP-Glucose and seven different concentrations of glucose-6-phosphate in the range of 0-3 mM. Data was fitted to a three parameter rectangular hyperbola $V=V_0+ a*x/(b+x)$ to calculate the AC_{50} for glucose-6-phosphate. The errors indicate the S.E.M. from three independent experiments

C. Crystallization and data collection of Gsy2p

It was observed that Gsy2p purified from fresh unfrozen cell pellets resulted in better diffraction quality crystals. Since 95% of the R580A/R581A/R583A crystals were physically twinned, the seeding method was used to improve the quality of the crystal. Microseeding of Gsy2p R580A/R581A/R583A mutant with the initial six-sided plate crystals resulted in 200 μm trapezoidal crystals that diffracted to about 3.0 Å. The triple mutant crystallized in two different crystal forms that belong to the same space group but differed slightly in the cell parameters (Table 2). The different crystal forms could not be differentiated visually and were identified during data collection and processing. The tantalum derivatives were prepared by soaking the triple mutant crystals. Upon incubation for six hours the tantalum bromide soaked crystals turned emerald green while the surrounding liquid became less colored and this property was used as an estimate for derivatization.

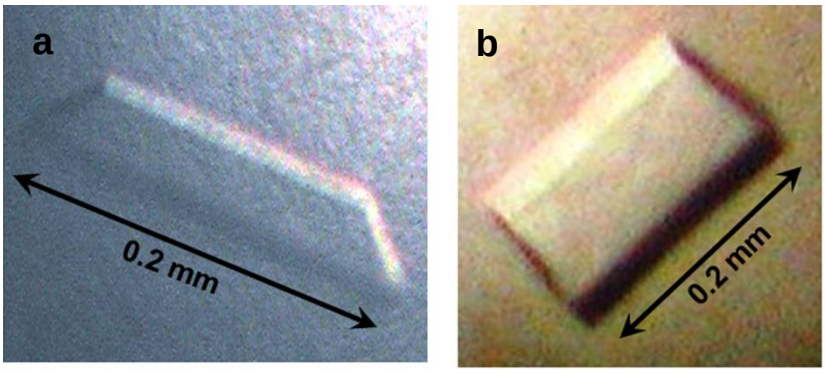

Figure 20. Gsy2p crystals.
(a) Primitive monoclinic crystal of Gsy2p R580A/R581A/R583A mutant in the $P2_1$ space group. (b) I centered orthorhombic crystals of of Gsy2p R589A/R592A mutant in the I222 space group.

The glucose-6-phosphate complex crystals of R589A/R592A mutant diffracted to 2.4 Å and crystals were formed only when the protein was co-crystallized with glucose-6-phosphate (minimum concentration of 25 mM). All the Gsy2p data sets were collected at the Synchrotron facility at the Advanced Photon source and the data collection statistics are summarized in Table 3.

Table 3. Data collection statistics

Data set	R580A3 Native1	R580A3 TaBr1	R580A3 TaBr2	R580A3 Native 2	R580A3 UDP	R580A3 Malto dextran	R589A2 G6P
Space group	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$	I222
Cell Dimensions							
a (Å)	96.5	97.3	96.5	93.9	96.6	96.6	192.7
b (Å)	166.1	167.6	167.3	161.8	167.2	166.7	206.9
c (Å)	121.0	122.2	121.6	121.6	121.2	121.1	205.8
α (°)	90.0	90.0	90.0	90.0	90.0	90.0	90.0
β (°)	103.4	103.3	103.1	103.1	102.7	103.2	90.0
γ (°)	90.0	90.0	90.0	90.0	90.0	90.0	90.0
Resolution (Å)	50-3.0 (3.1-3.0)	50.0 - 4.5 (4.6 -4.5)	50.0-4.5 (4.6-4.5)	50-2.9 (2.95-2.9)	50.0-3.5 (3.57-3.51)	50.0-3.0 (3.05-3.0)	50.0-2.4 (2.44-2.40)
R_{merge}(%)	8.2 (47.5)	10.9 (46.9)	12.5 (37.8)	8.7 (68.7)	11.0 (44.4)	8.4 (52.5)	8.7 (55)
I /σI	13.5 (2.1)	14.7 (1.7)	7.5 (1.9)	17.1 (2.3)	9.9 (1.7)	13.2 (2.0)	13.9 (2.0)
Completeness (%)	99.6 (96.8)	92.8 (94.7)	82.7 (64.1)	99.0 (98.7)	99.0 (92.7)	99.4 (94.7)	98.5 (79)
Redundancy	3.7 (2.6)	2.9 (2.6)	3.3 (3.4)	4.7 (4.5)	3.6 (2.7)	3.8 (3.1)	4.8 (3.1)

D. Structure solution of R580A/R581A/R583A Native1

1. Phasing multiple isomorphous replacement method

Gsy2p structure was solved using a native and two tantalum cluster derivative datasets of the R580A/R581A/R583A mutant. The derivatives diffracted to 4.5 Å and the Harker section of the isomorphous difference Patterson map generated with the native and derivative data sets showed clear peaks.

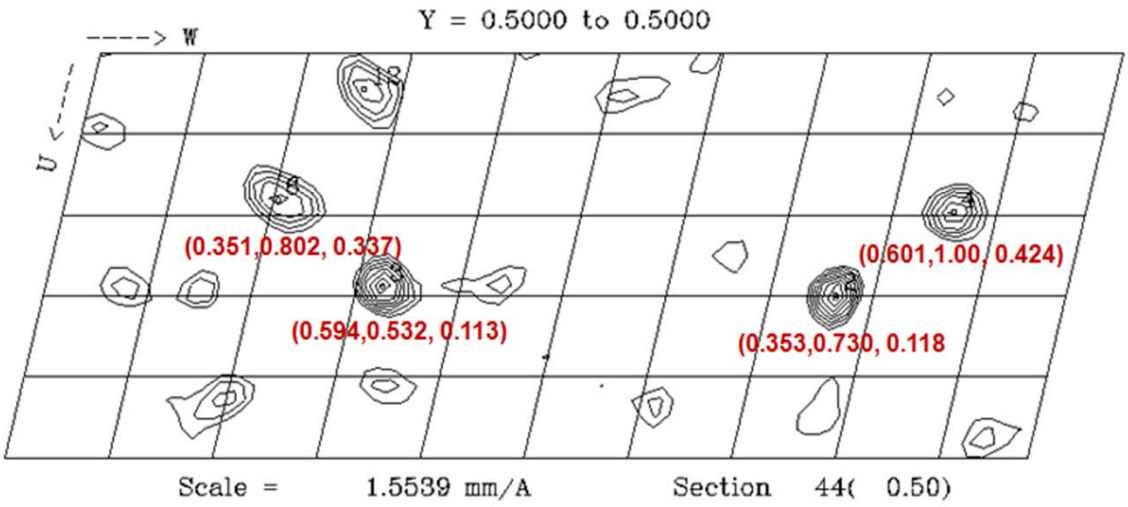

Figure 21. Isomorphous difference Patterson map.
Harker section of the Isomorphous Difference Patterson Maps at 5.5 Å generated with the Native1 and TaBr1 datasets. The x,y,z positions of the four common sites are marked in red.

The program SOLVE[87,88] was used to calculate the initial phases and the hexatantalum tetradecabromide compound was considered as a single tantalum site for the analysis. Although the derivatives diffracted to 4.5 Å, the isomorphous difference signal decreased significantly below 5.5 Å and hence the initial

phasing was done between 25 and 5.5 Å. Four common sites were identified in both the derivatives and a fifth site with lower occupancy was identified in one of the derivatives. The positions of the four common sites could be solved manually from the Harker peak positions. Together these derivatives provided phase information with a mean FOM of 0.47 and a Z-score of 43.8.

Table 4. Heavy atom solution from SOLVE

Site	Co-ordinate Position			TaBr1		TaBr2	
	X	Y	Z	Occup	H/Sig	Occup	H/Sig
1	0.601	1.000	0.424	0.86	9.8	0.84	10.0
2	0.353	0.730	0.118	1.04	12.0	0.88	12.0
3	0.351	0.802	0.337	0.98	12.4	0.94	13.7
4	0.594	0.532	0.113	0.75	10.0	0.75	10.1
5	0.799	0.565	0.519	0.20	2.8	-	-

The associated program RESOLVE identified a non-crystallographic 2-fold symmetry axis based on the tantalum sites and used the NCS averaging and density modification to provide an electron density map to 5.5 Å with a FOM of 0.74. These phases were sufficient to visually locate the biological tetramer within the asymmetric unit and the symmetry relationships amongst the 4 common tantalum sites in the two derivatives provided the location of the three 2-fold axes that define the tetramer.

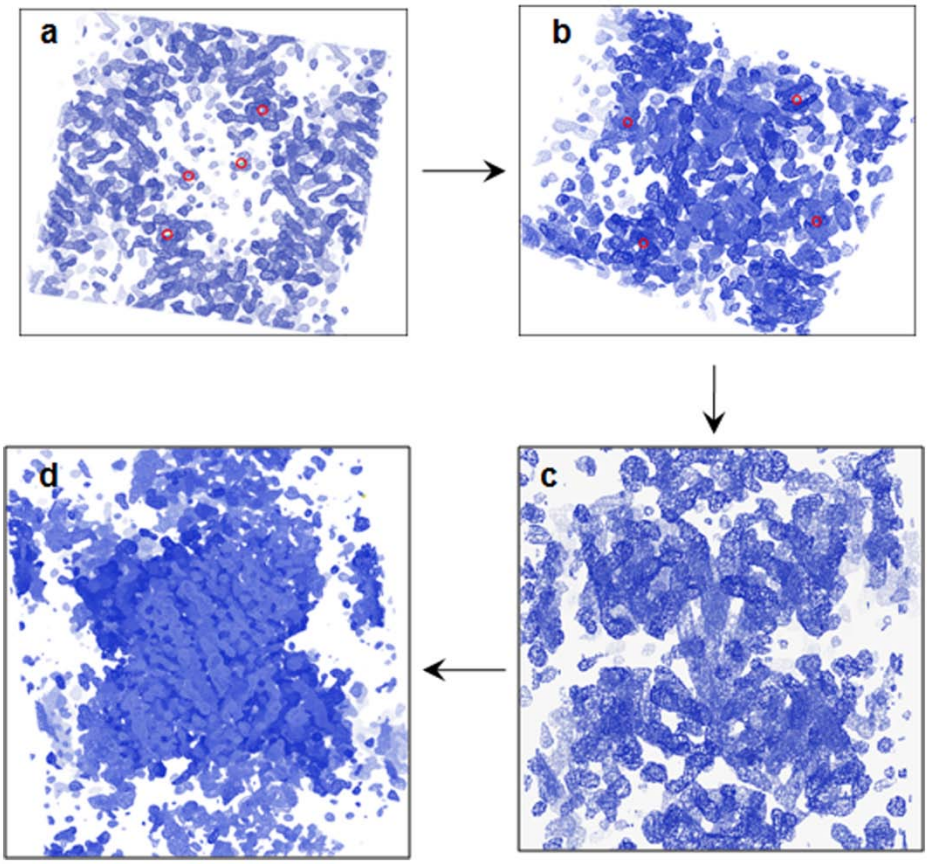

Figure 22. Gsy2p (R580A/R581A/R583A) phasing by multiple isomorphous replacement method.
(a) 2-fold averaged, 5.5 Å electron density map from RESOLVE. The four common tantalum cluster sites are highlighted. (b) 2-fold averaged, 5.5 Å electron density map from RESOLVE. The four common tantalum cluster sites in the tetramer highlighted. (c) 4-fold averaged, density modified 5.5 Å electron density map from RESOLVE. FOM 0.73 (d) 4-fold averaged, density modified, phase extended 3.0 Å electron density map from DM.

RESOLVE was rerun using explicit rotation matrices defining the 4-fold NCS relationships to provide phases to 5.5 Å with a mean FOM of 0.73. The electron density map produced using these phases clearly showed the presence of a tetramer in the asymmetric unit and showed good density for many α-helices but the β-sheets were less well defined. Phase extension to 3.0 Å was

accomplished using the program DM as implemented in the CCP4 suite[89]. This phase extension protocol utilized a total of 250 steps from 6.0 to 3.0 Å and the same 4-fold averaging matrices utilized by RESOLVE. The resulting 3.0 Å electron density map was marginally interpretable, but the general relationship of the protein fold to the bacterial glycogen synthase enzymes could be identified.

2. Phase extension by phase combination approach

To improve the map quality, we used the phase combination approach implemented in the program Phenix[90,91] ("maps_only" subroutine). A partial model of the yeast Gsy2p monomer was generated by docking elements of the secondary structure from a polyalanine model of the Agrobacterium GS monomer (PDB code 1RZU) into a single subunit within the 3.0 Å electron density map. The tetramer was generated by applying the NCS operators to this partial monomer structure. The phase information from this partial model and the experimental phase information from SOLVE were combined to extend the usable phases from 5.5 to 3.0 Å. Two iterations of phase combination with increasingly larger partial models within the program Phenix produced phases that yielded a completely interpretable 3.0 Å electron density map, from which a poly-alanine model of the protein was completely rebuilt using the graphics program COOT[92]. A structure based sequence alignment of the glycogen synthases was used to assist the sequence assignment to the initial poly-alanine model.

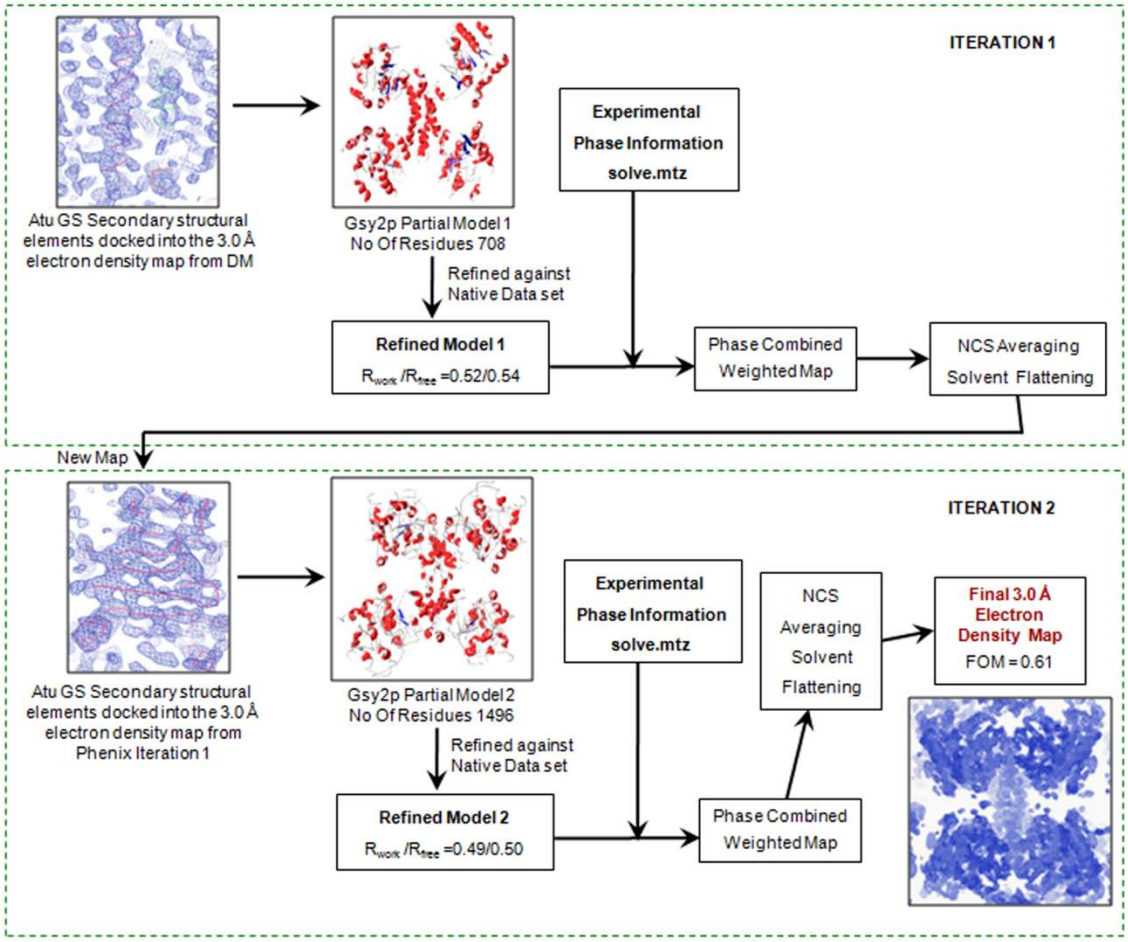

Figure 23. Schematic representation of the phase extension approach. The current electron density maps from DM and RESOLVE were used to generate a partial model of Gsy2p, which was refined against the native data set. The phase information from this refined model and the experimental phase information (isomorphous heavy metal signal from SOLVE) were combined to give a phase combined map, which was further subjected to 4-fold NCS averaging and solvent flattening to generate the new electron density map. Two iterations of the phase combination approach resulted in a completely interpretable 3.0 Å electron density map.

E. Refinement of Gsy2p structures

All the refinement protocols utilized NCS restraints for each subunit or each domain of the subunits. The two R580A/R581A/R583A structures were refined to 3.0 and 2.9 Å and final R_{work} / R_{free} values of 20.3/25.4 and 22.1/26.1. The glucose-6-phosphate bound activated conformation of the R589A/R592A mutant structure was refined to R_{free} = 24.3% and R_{work} = 20.3% (Table 5) . The asymmetric unit of both the crystals contain a tetramer. The triple mutant structures contain residues 2 to 639 with three major breaks, where the density was not sufficient to build the model. These disordered regions showed clear density in the glucose-6-phosphate activated conformation. In addition, chains B and C in the activated conformation structure exhibited clear electron density for an additional seven residues at the C-terminus. All the final models displayed good stereochemistry as analyzed by the program PROCHECK[93].

Table 5. Data collection statistics

Data	R580A3 Native1	R580A3 Native 2	R580A3 UDP	R580A3 Malto dextran	R589A2 G6P
Resolution (Å)	3.0	2.9	3.5	3.0	2.4
No. reflections	70557	80301	46330	73079	148274
R_{work} / R_{free}(%)	20.3/25.4	22.1/26.1	21.4/24.4	20.9/24.7	20.3/24.3
No of atoms					
Protein	19770	19703	19707	19742	20818
Ligand/ion	80	75	160	260	124
B-factors					
Protein	64.37	70.29	103.90	82.10	57.41
Ligand/ion	90.05	90.50	130.72	112.14	47.11
R.m.s deviations					
Bond lengths (Å)	0.008	0.009	0.009	0.008	0.009
Bond angles (°)	1.07	1.17	1.12	1.15	1.15

F. Structure of Gsy2p R580A/R581A/R583A

The asymmetric unit of the native 1 R580A/R581A/R583A crystals contains a single tetramer and the final refined protein model contains residues 2-205, 208-277, 285-401, 414-540 and 546-639. The C-terminal 65 residues, which contain the regulatory phosphorylation sites, are not observed in the structure. SDS-PAGE analysis of the crystals reveals that less than 10% of the protein in the crystal has been proteolyzed. This suggests that the C-terminal region of the yeast Gsy2p is crystallographically disordered, rather than selectively proteolyzed during crystallization. This is consistent with predictions of intrinsic disorder by the program PONDR®[94] that identifies the final 80 residues as having a high probability of disorder.

1. Overall fold and oligomeric arrangement

Like other GTB family members, the yeast Gsy2p monomer folds into two structural domains dominated by their individual Rossmann-folds, each with a central six-stranded parallel β–sheet flanked on either side by α-helices. The N-terminal Rossmann domain has two inserts, one between the strands β2 and β5, and the other between β7 strand and α10 helix (Figure 24). The polypeptide crosses between the two domains with the linker peptide consisting of strand β9, and helices α12 and α13, which together form one of the two inter-domain connections. The second inter-domain connection is made upon exit of the C-terminal Rossmann domain where the two long helices (α22 and α23) form a

characteristic feature of GT-B glycosyl transferases with these inter-domain helices. The C-terminal Rossmann domain of all eukaryotic synthases has a unique insert between strand β11 and helix α17 (Figure 24). This insert's major structural feature is a pair of long helices that extend away from the Rossmann-fold and form the majority of the inter-subunit interface. The cluster of conserved arginine residues found in eukaryotic synthase enzymes is located in the first of the two inter-domain helices (α22), which we will refer to as the regulatory helix, in consideration of the regulatory nature of these arginine residues. Lastly, a small, unidentified, fragment of polypeptide between 4-6 residues long is present within the substrate-binding cleft of monomers C and D. These could be either residues comprising the N-terminal His-Tag extension from the symmetry related monomers A and B which are positioned 18Å away, or a short region of the disordered C-terminal extension. There is insufficient electron density for unambiguous sequence assignment of these residues, hence they have been modeled as a stretch of poly-alanine residues.

The most striking feature of the yeast Gsy2p is the tetrameric arrangement of the subunits, and relatively limited intersubunit contact surface between the two dimers that comprise the enzyme's dimer-of-dimers arrangement. The primary subunit interface is formed by the coiled-coil helical arrangement of helices α15 and α16, which represent the unique insertion in the C-terminal Rossmann domain (residues 365-431). A 12 amino acid loop between these two helices (residues 401-412) is disordered and not visible in the R580A/R581A/R583A structures.

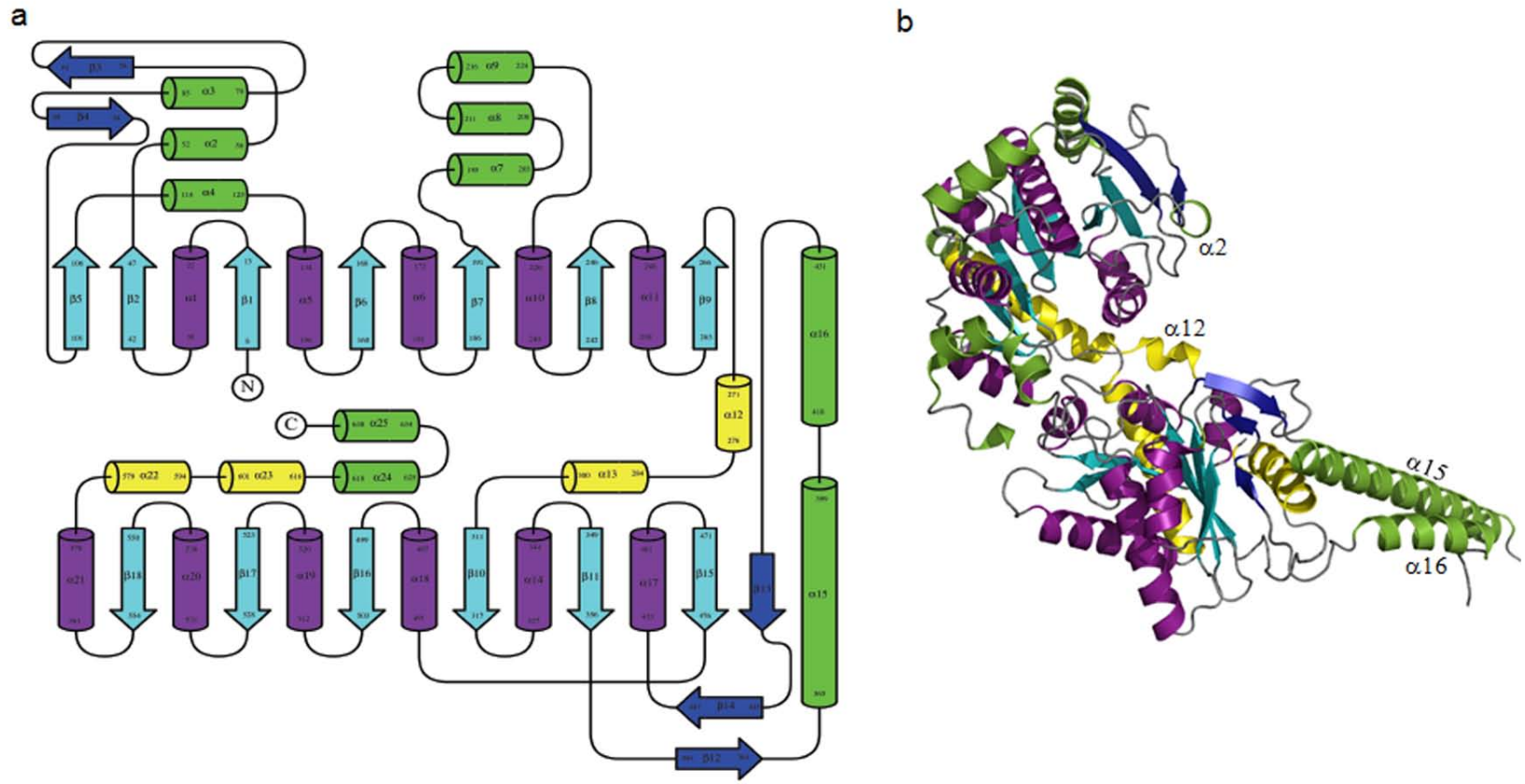

Figure 24. Structure of Gsy2p monomer.
(a) Topology diagram of the protein fold of Gsy2p. (b) Ribbon representation of the structure of Gsy2p monomer. The conserved secondary structural elements of the core Rossmann domains are represented in purple and cyan, the linker region in yellow and the unique eukaryotic inserts in blue and green.

The major interaction between dimer pairs is made by the C-terminal region of helix α15 and the loop region that connects the linker peptide helix α15 with the β10 strand of C-terminal Rossmann domain (Figure 25). The dimer interace is formed by the association of helix α16 of one monomer with helix α2 and the loop connecting the strands β4 and β5 of the other monomer.

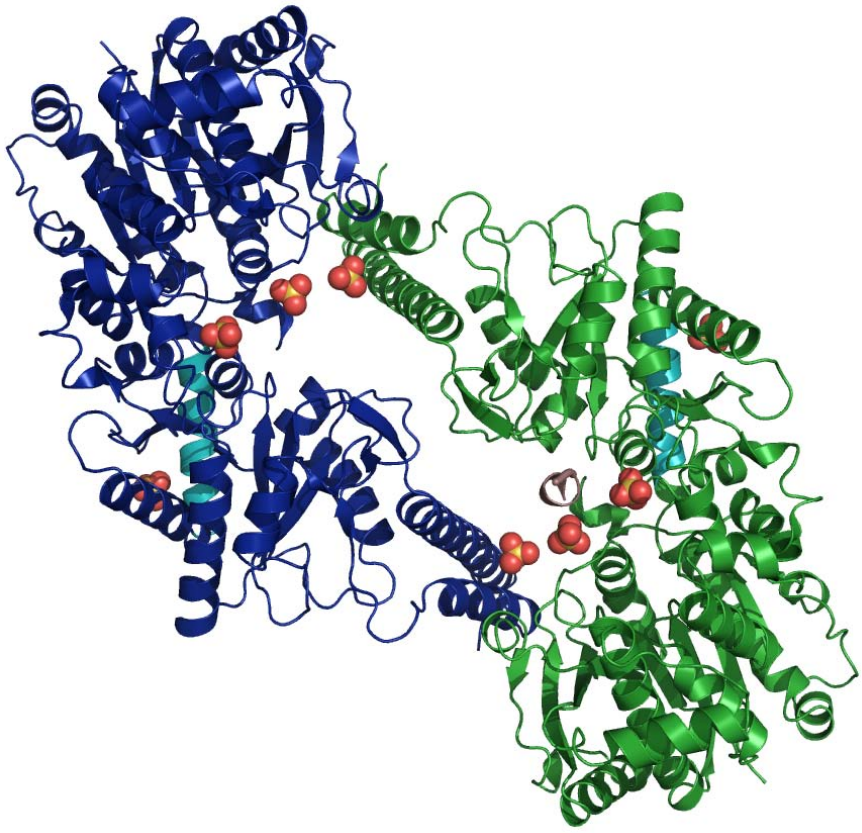

Figure 25. Dimer arrangement of Gsy2p.
Ribbon representation of the structure of Gsy2p dimer with the bound sulfates represented as space filling models. Different colors represent individual subunits and the regulatory helix is shown in cyan.

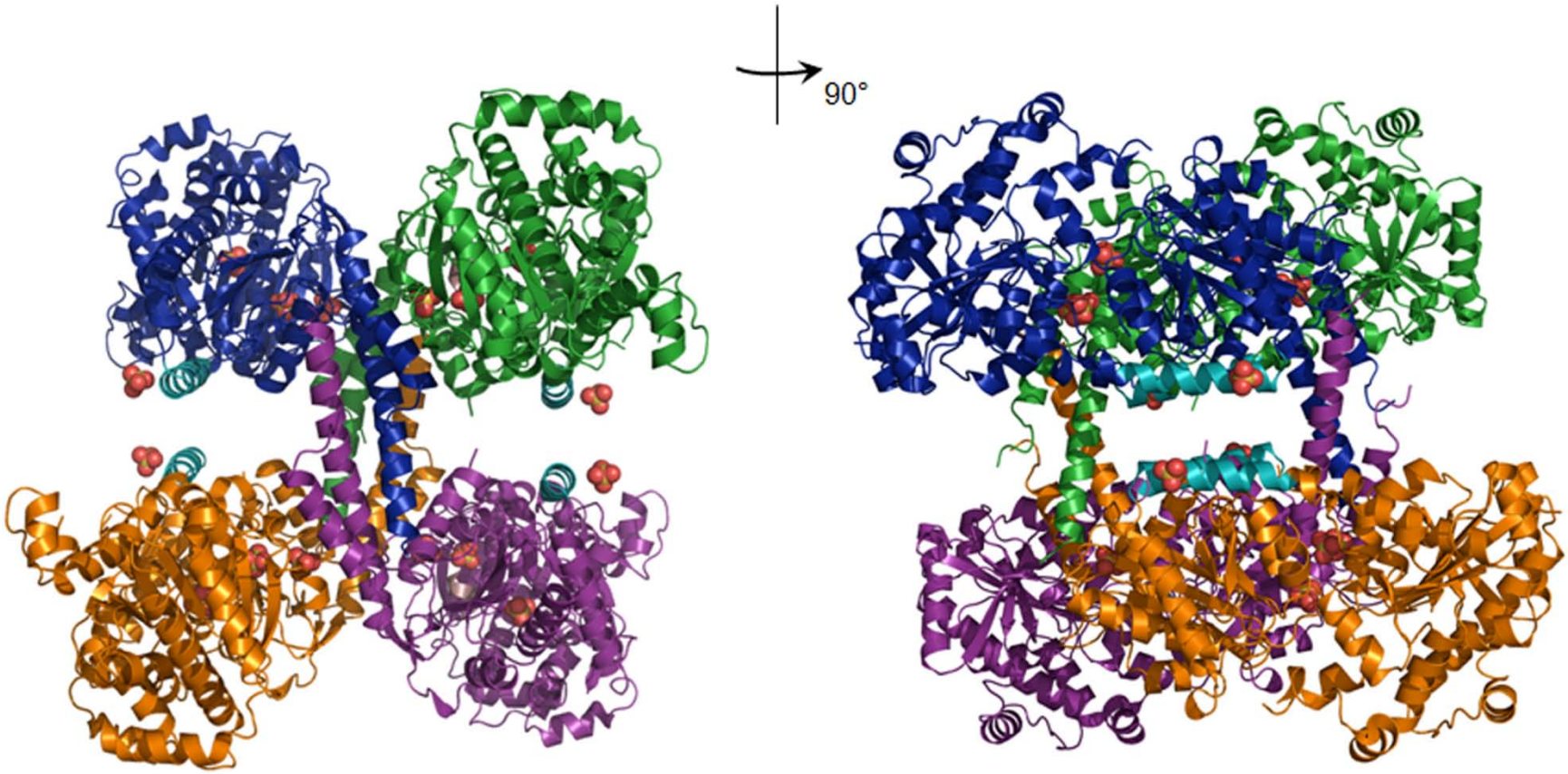

Figure 26. Tetramer arrangement of GSy2p.
Ribbon representation of Gsy2p R580A/R581A/R583A tetramer with the bound sulfates in space filling models. Different colors are used to represent different subunits and the regulatory helices are shown in cyan.

2. Arginine cluster

The regulatory helices ($\alpha 22$) containing the arginine cluster are located across from one another along opposing sides of one of the molecular 2-fold axes in the tetramer and the orientation of the helical axis is orthogonal to the helical bundles forming the tetrameric interface (Figure 26). A surface pocket is present in the N-terminal region of the regulatory helix in the proximity of the side chains of the residues 583 and 587 (3^{rd} and 4^{th} arginines of the conserved cluster), which are pointing towards this cavity (Figure 27b). Due to the position of these key arginines and the mixed hydrophilic and hydrophobic nature of the cleft, we hypothesized that this cleft forms the glucose-6-phosphate binding pocket and these two arginines are likely involved in hydrogen bonding or electrostatic interactions with the phosphate group of glucose-6-phosphate.

To test this hypothesis we reintroduced Arg 583 and Arg 587 into their respective triple arginine mutants, forming the R580A/R582A and R588A/R592A double mutants, and tested their glucose-6-phosphate sensitivity. Reintroduction of either arginine into their respective triple mutants restored glucose-6-phosphate sensitivity (Table 6).The R588A/R592A double mutant displayed specific activities and activity states similar to those of the wild-type phosphorylated enzymes (Table 6).

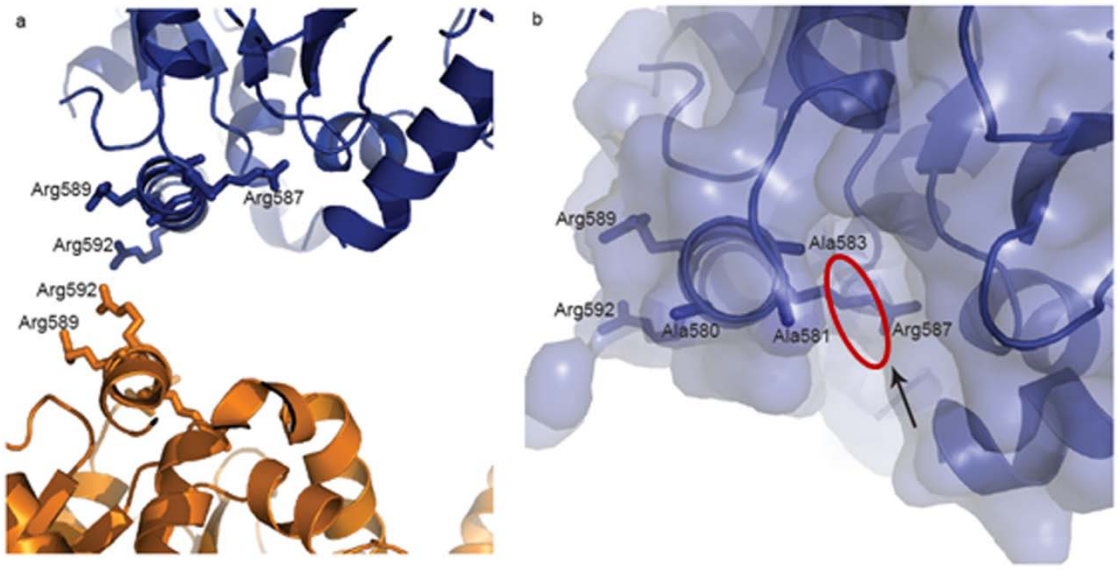

Figure 27. Arginine cluster and glucose-6-phosphate binding pocket of Gsy2p.
(a) Ribbon diagram of Gsy2p highlighting the regulatory helices at the dimer-dimer interface. (b) Ribbon diagram and molecular surface representation of the Gsy2p monomer highlighting the glucose-6-phosphate binding pocket (black arrow). The three alanines and arginines at the interface are labeled.

To study the role and influence of the arginine residues in phosphorylation mediated inhibition, we used the peptide ligation approach to create semisynthetic phosphorylated proteins of the mutants and tested their activity. Both the triple arginine mutants were nearly insensitive to inhibition by phosphorylated Thr668 and exhibit similar activities after ligation to both the 35mer and 49mer phosphopeptides. Interestingly, the activity ratios of the peptide-ligated constructs were inversely correlated with the number of arginines residues in this first cluster, with the R580A/R581A mutant having the highest activity ratio and the wild-type sequence the lowest activity ratio.

Table 6. Specific activity and activity ratio of Gsy2p arginine mutants

Enzyme	Synthase Specific activity (mmol. min^{-1})		Activity Ratio -G6P/+G6P
	No G6P	7.2mM G6P	
Wild Type	1.03 ± 0.03	1.73 ± 0.03	0.60 ± 0.01
R580A	0.99 ± 0.01	1.55 ± 0.03	0.64 ± 0.01
R580/R581A	1.07 ± 0.05	1.59 ± 0.05	0.67 ± 0.02
R580A/R581A/R583A	0.55 ± 0.02	0.51 ± 0.03	1.09 ± 0.02
R587A/R589A/R592A3	0.66 ± 0.01	0.58 ± 0.02	1.14 ± 0.05
R589/R592A	0.17 ± 0.01	1.56 ± 0.07	0.11± 0.01
Wild type-35mer Phosphopeptide ligated	0.27 ± 0.02	1.96 ± 0.04	0.14 ± 0.01
R580A/R581A/R583A -35mer Phosphopeptide ligated	0.47 ± 0.01	0.44 ± 0.01	1.05 ± 0.02
R587A/R589A/R592A-35mer Phosphopeptide ligated	0.41 ± 0.01	0.39 ± 0.02	1.04 ± 0.05
Wild type-49mer Phosphopeptide ligated	0.10 ± 0.01	1.92 ± 0.04	0.05 ± 0.01
R580A-49mer Phosphopeptide ligated	0.34 ± 0.02	1.74 ± 0.10	0.20 ± 0.01
R580A/R582A-49mer Phosphopeptide ligated	0.58 ± 0.01	1.50 ± 0.01	0.39 ± 0.01
R580A/R581A/R583A -49mer Phosphopeptide ligated	0.43 ± 0.01	0.42 ± 0.01	1.04 ± 0.04
R587A/R589A/R592A-49mer Phosphopeptide ligated	0.38 ± 0.01	0.35 ± 0.01	1.09 ± 0.09

GS activity was measured in the presence of 6.7mg/ml of glycogen and 4.4 mM UDP-glucose in the absence and presence of 7.2 mM glucose-6-phosphate. The errors indicate the S.E.M. from three independent experiments

G. Allosteric activation of Gsy2p – Structure of R589A/R592A

1. Glucose-6-phosphate binding

The activated conformation of Gsy2p was obtained by co-crystallizing the R589A/R592A mutant with glucose-6-phosphate. Glucose-6-phosphate is bound in all the four subunits of the tetramer in the putative glucose-6-phosphate binding cleft located between the regulatory helix (α22) and helix α13. The side chains of residues His 286, Lys 290, His 500, Arg 583, Arg 587 and Arg 580 form hydrogen bonding interactions with the 6-phosphate group of the activator (Figure 28). The glucose ring is positioned in such a way that the C1’ hydroxyl group forms a hydrogen bond with the side chain of Gln 283 from one subunit and the C2’ hydroxyl group hydrogen bonds with the His280 from the opposing subunit. In addition, the C6’ and O5 oxygen atoms of the glucose ring interact with the side chain of Arg 587. The C2’ hydroxyl is also an average of 3.7 Å from the side chain of Asn 284. The faces of the glucose ring are held between the side chains of Ala 287 and Ile 584.

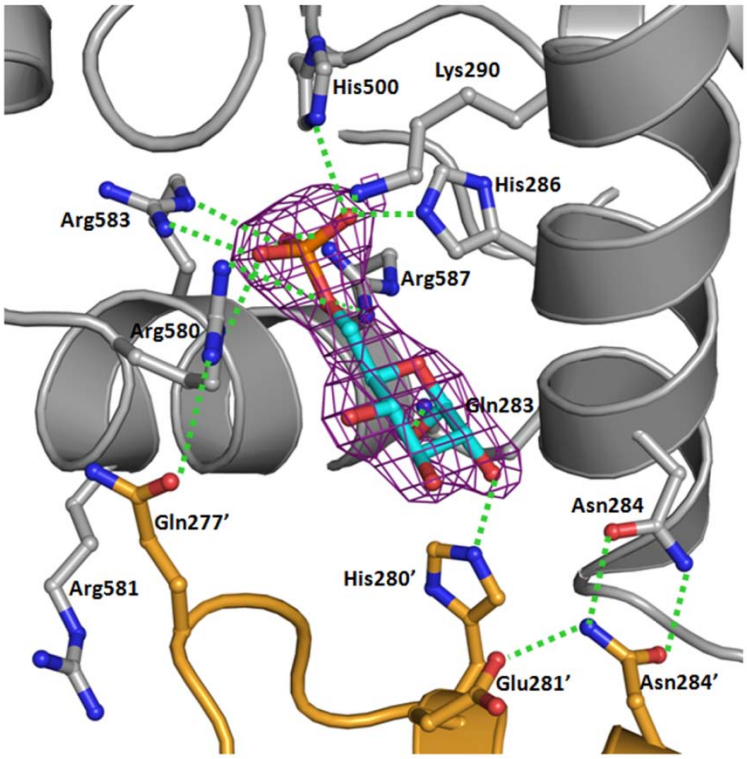

Figure 28. Glucose-6-phosphate binding in Gsy2p.
Ribbon representation of Gsy2p with the bound glucose-6-phosphate and the interacting residues in ball-and-stick model. Different colors are used for representing different subunits. The map shown is the original 2fo-fc electron density map for the bound glucose-6-phosphate (contoured at 1 s.d of the map). The interacting residues are labeled (the prime sign (') indicates the residues from second subunit) and the hydrogen bonding interactions are represented as dashed green lines.

2. Conformational changes induced by Glucose-6-phosphate

Binding of glucose-6-phosphate induces an ordering of the structure for the loop connecting helix α13 to the first strand (β10) of the C-terminal Rossmann domain – residues 277-284 (Figure 29b). The ordered loop stabilizes the activated conformation by interacting with the glucose-6-phosphate and side chain residues of the partner subunit and increasing the relative distances of the regulatory helices from each other by 3.4 Å. The major interactions stabilizing

this interface include the aforementioned interactions between His 280 and the C2' hydroxyl of glucose-6-phosphate, hydrogen bonds between the equivalent side chains of Asn284 and an interaction between Arg 581 and the main chain carbonyl oxygen atoms of residues 276 and 277.

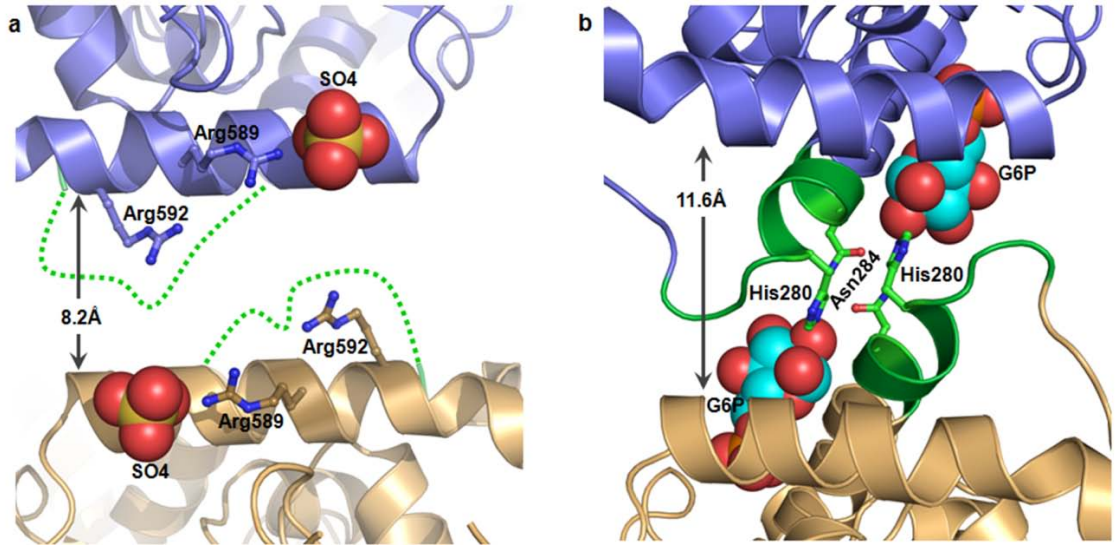

Figure 29. Glucose-6-phosphate binding induced conformational change at the dimer interface.
Ribbon representation of the regulatory helix dimer interface of Gsy2p with the bound sulfate (panel a – R580A/R581A/R583A) and glucose-6-phosphate (panel b – R589A/R592A) in space fill models. The region between residues 277 and 285 that is ordered upon glucose-6-phosphate binding is shown in green.

To analyze the conformational transition induced by glucose-6- phosphate binding, we used the native-1 structure of the R580A/R581A/R583A Gsy2p as the basis upon which to describe the changes associated with the activated conformation of the R589A/R592A double mutant. Due to the complexity of the motions, the analysis will divide each subunit into three structural domains – the

N-terminal Rossmann domain (residues 2-272), the C-terminal Rossmann domain (residues 312-357, 453-577) and the long intersubunit helices (residues 365-399, 416-434).

When the two structures are superimposed using only the intersubunit helices for the AC dimer pair, the BD dimer pairs in the two structures are displaced by an overall translation of 14.2 Å combined with a 7° rotation (Figure 30).

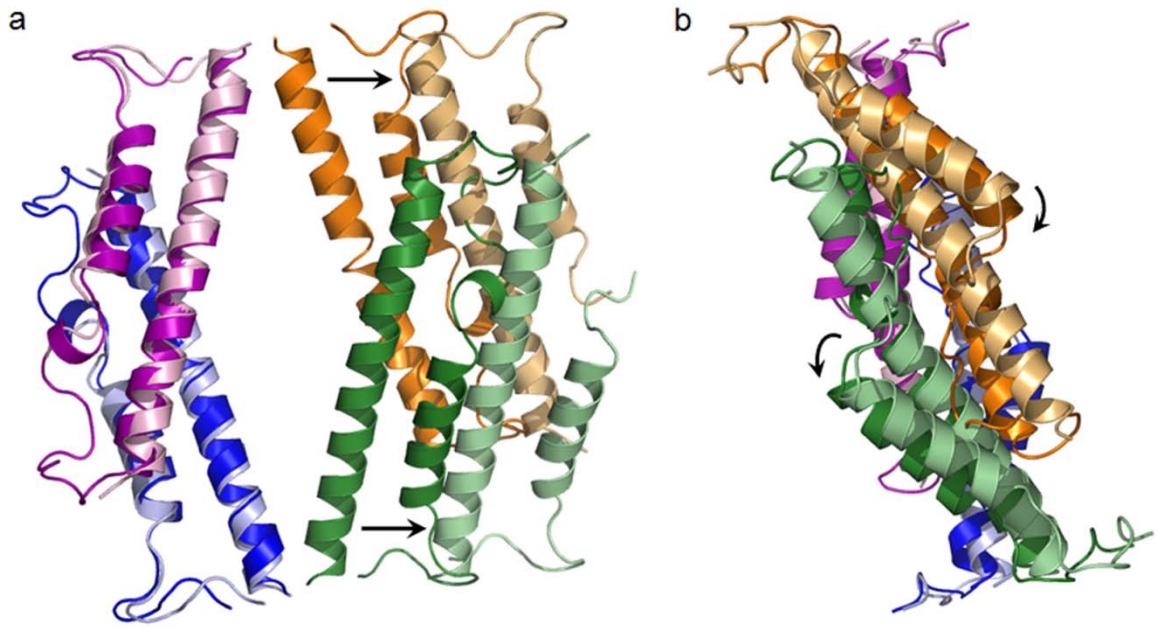

Figure 30. Glucose-6-phosphate induced rotational and translational motions.
Ribbon representation of the aligned long helices of Gsy2p in the two conformations highlighting the glucose-6-phosphate induced translational (a) and rotational movements (b) Alignment was generated by superimposing the long helices of the a and C subunits using the LSQ-superpose program. Different colors are used to represent different monomers and the light and dark shades are used to represent the R580A/R581A/R583A and R589A/R592A structures respectively. The translational and rotational movements are marked by black arrows.

Within each monomer, the binding of the activator induces a rotation of 8.5° of the C-terminal Rossmann domain when the intersubunit helices are aligned to form a basis for comparison (Figure31). In addition to the rotation of the C-terminal domain, the N-terminal Rossmann domain is also rotated away from the central interface by an additional 4.5° (Figure 31).

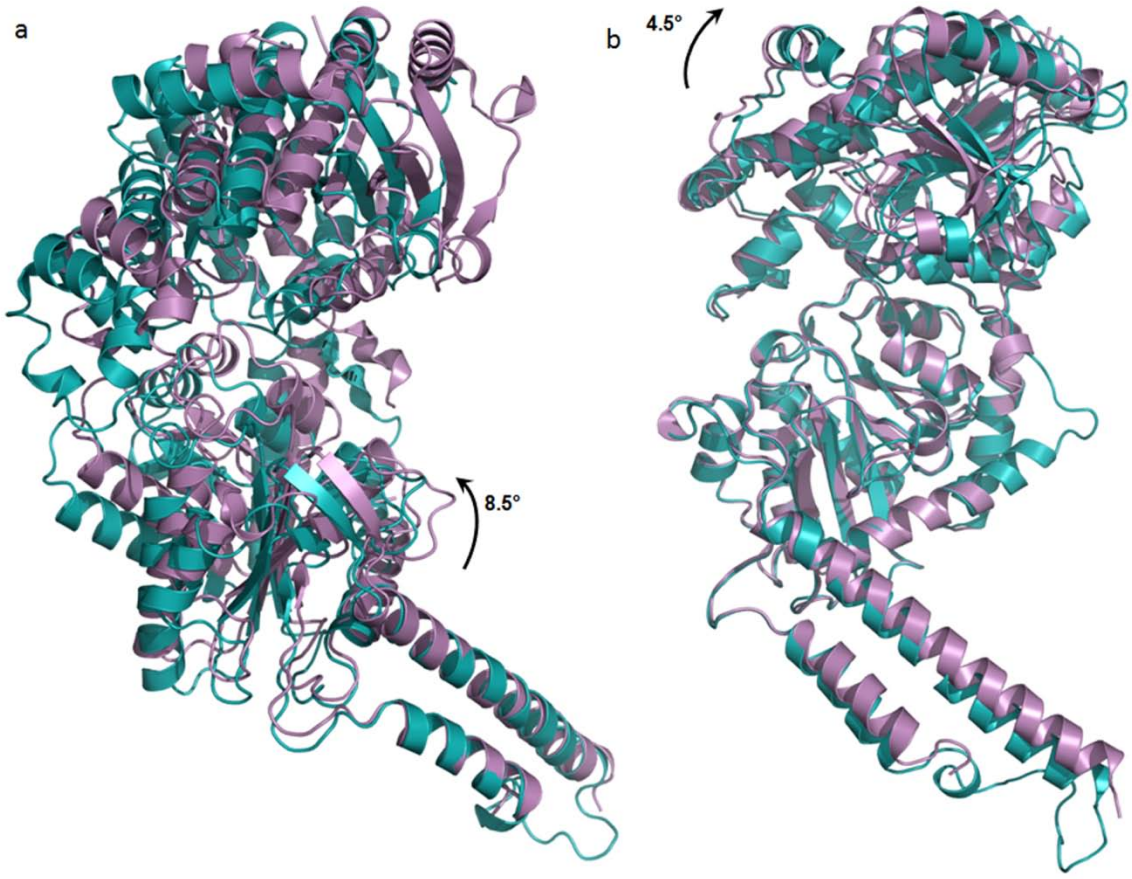

Figure 31. Glucose-6-phosphate binding induced conformational change in the Gsy2p monomer.
Ribbon representation of the superposed basal state and activated state monomers. Different colors represent different conformations (violet – R580A/R581A/R583A and Teal – R589A/R592A). (a) The alignment was generated by superposing the long helices of chain A (residues 365-399 and 416-434) of both the conformations using the superpose program of the CCP4 suite. (b) Alignment was generated by superposing the C-terminal Rossmann domain of R580A/R581A/R583A chain A (residues 312-357 and 453-577) with the long helix aligned chain A of R589A/R592A.

In additional to the rotation of the Rossmann domains, the other major change observed at the dimer (AD) interface is in the positions of helix α16 and the loop connecting strand β16 and helix α18 (loop β16-α18) . In the R580A/R581A/R583A structure, the β16- α18 loops are located at one of the molecular 2-folds and contact each other, while the α16 helix forms interactions with strand β3 and helix α2 within the unique sequence insertion found in the N-terminal Rossmann domain. Activation by glucose-6-phosphate induces a translation of the α15 and α16 helices toward the center of this interface, moving the β16- α18 loops away from the local 2-fold axis and forming new interactions with helices α15 and α16 from the opposing subunit. This translation, coupled with the rigid-body rotations of the subunits away from the central interface disrupts the interactions between α16 and the N-terminal domain, thereby releasing the N-terminal domain and active site from constraining interactions with the C-terminal domain of its partner subunit (Figure 32).

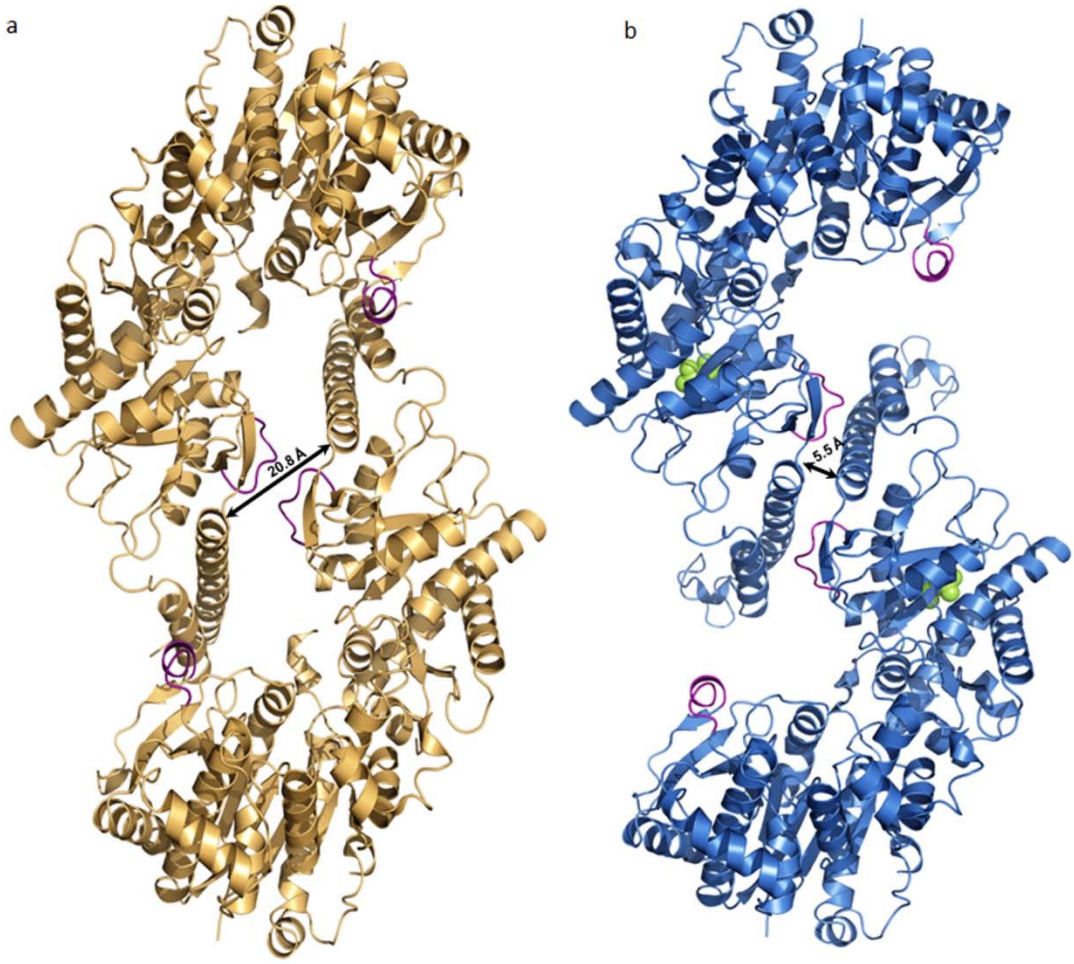

Figure 32. Glucose-6-phosphate binding induced conformational change in the AD dimer.
Ribbon representation of the AD dimer pair in the basal state and activated state (b) with the bound glucose-6-phosphate in space filling model. The β16-α18 and β2-α2 loops are highlighted in purple.

H. Substrate Binding in Gsy2p

1. UDP-binding pocket

UDP binds to Gsy2p- R580A/R581A/R583A within the large inter-domain cleft that separates the two Rossmann-fold domains and interacts with residues within the C-terminal domain. The UDP binding pocket is defined by helices $\alpha 18$, $\alpha 19$ and the loops between $\beta 16–\alpha 19$, $\beta 15–\alpha 18$ and $\beta 10–\alpha 14$. The uridine ring is stacked between the aromatic side chains of Phe 480 and Tyr 492 and the 4-oxo group of the uridine ring is positioned to form a hydrogen bond with the main chain nitrogen atom of Leu481. The ribose ring is positioned within hydrogen-bonding distance to the conserved residue Glu 517 such that the side chain appears to interact with both the 2' and 3' hydroxyls of the ring. The phosphate groups of the UDP appear to interact with residues Arg 320, Lys 326, Thr 514 and the main-chain atoms of residues Tyr 513 and Thr 514. The strictly conserved residue Glu 509, which is proposed to be the catalytic nucleophile, is also in close proximity to the bound UDP molecule.

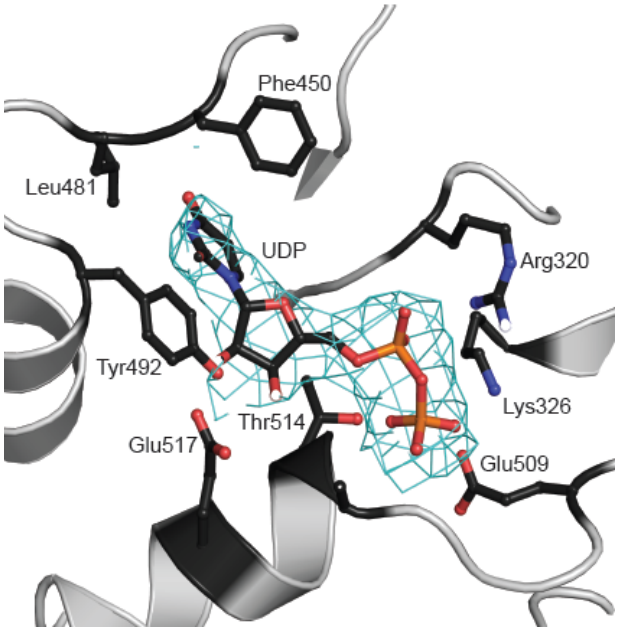

Figure 33. UDP binding in Gsy2p.
Ribbon representation of the Gsy2p UDP binding pocket with the bond UDP and interacting residues in ball and stick model. The map shown is the original $2F_o$ - F_c electron density map for the bound UDP (contoured at 1 s.d of the map). The residues involved in binding are labeled.

2. Maltodextran binding pocket

The structure of the Gsy2p complex with maltodextran was solved in the $P2_1$ space group to a resolution of 3 Å. We observe two binding sites for maltodextrans in each subunit – one on the surface of the N-terminal domain (site 1) and the other (site 2) is located in a crevice near the dimer interface in the C-terminal domain that defines the structural boundary between the conserved catalytic core of the enzyme and the unique sequence insertion that forms the oligomerization interface (Figure 40c). In each of the two maltodextran binding sites, we observe four ordered glucose residues bound within depressions on the surface of the enzyme that are lined with residues that are conserved amongst eukaryotic GS enzymes.

Helix $\alpha 4$ is located adjacent to the 1' end of the maltodextran moiety in the N-terminal domain of Gsy2p. The $\alpha 4$ helix is a unique secondary structural element present in all the eukaryotic enzymes and is involved in defining one side of site 1 in Gsy2p. The remainder of the site is contributed by residues within helix $\alpha 5$ which forms the base of the maltodextran site and runs more or less parallel with the carbohydrate moiety and the helix $\alpha 6$. Residues Trp 118, Tyr 145, Trp 149 and Ile184 form the hydrophobic pocket in which the glycan residues reside. Hydrogen bonding interactions between the glucose polymer and the enzyme are mediated by the side chains of Glu 117, Asp 121, Glu 153, Arg182 and Arg 86. The site 1 region on the C and D subunits are utilized in

crystal packing contacts and are unavailable to interact with maltodextran in this lattice.

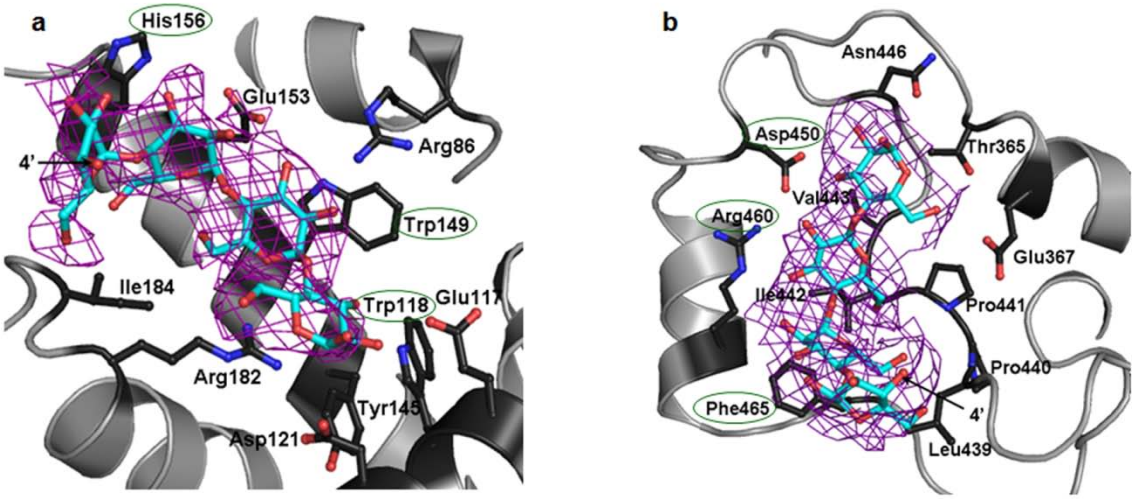

Figure 34. Maltodextran binding sites in Gsy2p.
Ribbon representation of the Gsy2p Maltodextran binding site 1(a) and site 2(b) with the bond Maltose polymer and interacting residues in ball and stick model. The map shown is the original $2f_o$ - f_c electron density map for the bound UDP (contoured at 1 s.d of the map). The residues involved in binding are labeled. The residues selected for mutation are circles in green.

The second maltodextran binding site is located in the C-terminal domain at the structural boundary between the long helices that form the oligomeric interface and the C-terminal Rossmann-domain. The orientation (1' to 4') of the maltodextran moiety is essentially anti-parallel to the direction of the beta-strand lying on top of the C-terminal Rossmann-domain. The inner surface of this site is lined by a hydrophobic stretch of residues between Leu 439 and Val 443, as well as Phe 465. This hydrophobic pocket is surrounded by residues Thr 365, Glu 367, Asn 446, Asp 450 and Arg 460 which provide hydrogen bonding interactions with the bound glucose polymer.

To test the influence of these maltodextran binding sites on Gsy2p activity, we created two triple alanine mutants; W118A/W149A/H156A for site 1 and D450A/R460A/F465A for site 2 and compared the kinetic properties of the mutants with the wild-type Gsy2p (Figure 35 and Table 7). When compared to the wild-type Gsy2p the specific activity of the mutants decreased by approximately three to five fold in the presence or absence of glucose-6-phosphate. Kinetic analysis with varying concentrations of glycogen showed that in the absence of glucose-6- phosphate, the V_{max} of the mutants decreased five times when compared to the wild type. With saturating glucose-6-phosphate, this decrease was comparatively less - about 1.5 times. In the absence of glucose-6-phosphate, the $S_{0.5}$ for glycogen increased from 0.25 mg/ml for the wild type to about 2 mg/ml for the mutants. Interestingly, in the presence of glucose-6-phosphate, the $S_{0.5}$ for glycogen was decreased for the wild type, whereas it was increased with the mutant enzymes (Table 7).

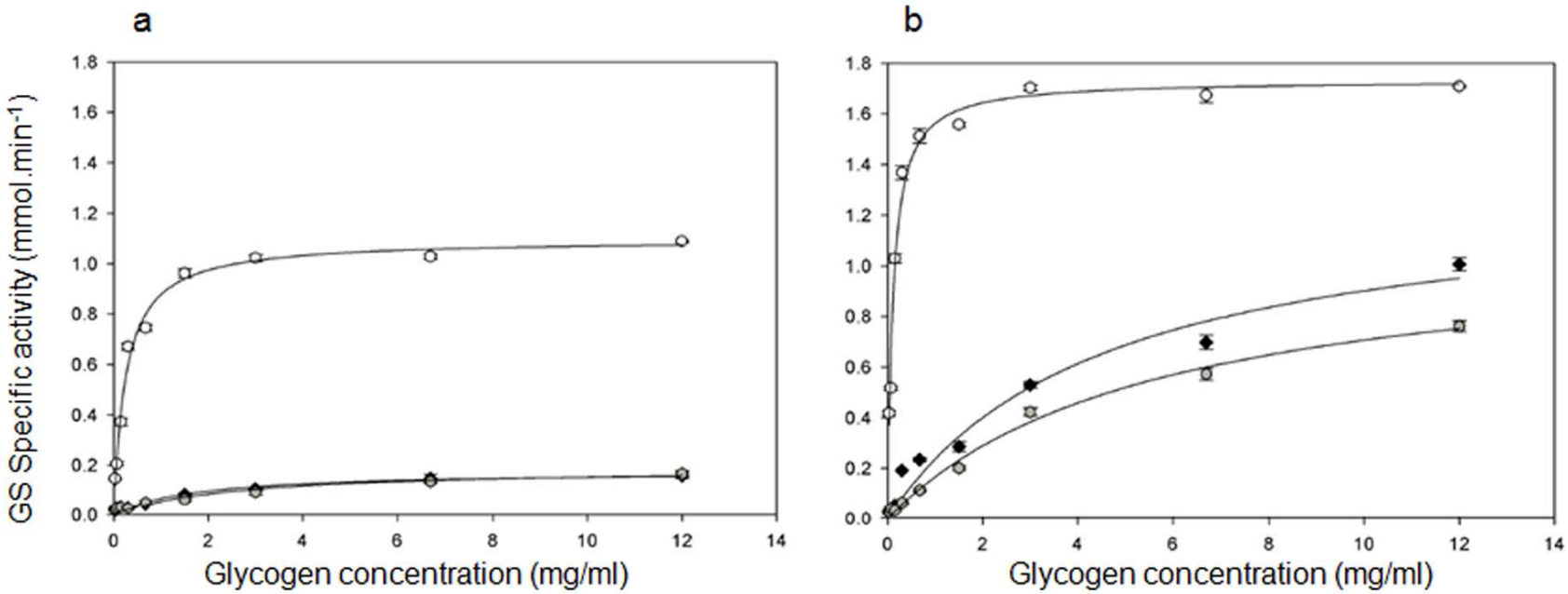

Figure 35. Glycogen titration of Gsy2p.
The specific activity of wild-type (white circle), W118/W149/H156A3 (black diamond) and D450/R460/F465A3 (grey hexagon) enzymes were measured in the absence (a) and presence (b) of saturating 7.2 mM glucose-6-phosphate and varying concentrations of glycogen between 0.03 – 12 mg/ml. The activity measurements were done with 4.4 mM UDP-glucose. The error bars represent the standard error of mean from three independent experiments.

Table 7. Specific activity and kinetic parameters of Gsy2p

Enzyme	Specific activity (mmol. min^{-1})		V_{max} (Glycogen Varied)		$S_{0.5}$ Glycogen	
	No G6P	7.2mM G6P	No G6P	7.2mM G6P	No G6P	7.2mM G6P
Wild Type	1.03 ± 0.03	1.73 ± 0.03	1.10 ± 0.01	1.72 ± 0.01	0.25 ± 0.01	0.11 ± 0.01
118/149/156A3	0.18 ± 0.01	0.56 ± 0.04	0.18 ± 0.02	1.31 ± 0.03	1.72 ± 0.32	4.58 ± 0.25
450/460/465A3	0.14 ± 0.01	0.49 ± 0.01	0.20 ± 0.02	1.11 ± 0.03	2.72 ± 0.77	5.65 ± 0.06
118/149/156/ 450/460/465A6	0.03 ± 0.00	0.03 ± 0.00	ND	ND	ND	ND

ND – Not Determined

Specific activity was measured in the presence of 6.7mg/ml of glycogen and 4.4 mM UDP-glucose in the absence and presence of 7.2 mM glucose-6-phosphate. Glycogen titration was performed under the same conditions of UDP-glucose and glucose-6-phosphate with nine different concentrations of glycogen in the range of 0.03 to 12.0 mg/ml. The errors indicate the S.E.M. from three independent experiments

I. Insight into inhibition by phosphorylation

1. Sulfate as phosphomimetic in R580A/R581A/R583A structure

There are four sulfate ions bound to each of the four monomers in the native1 R580A/R581A/R583A structure. Within each monomer, sulfate-1 is bound at the position equivalent to that of the distal phosphate of the UDP molecule and interacts with the side chains of Arg 320 and Lys 326, as well as the main chain atoms of residues Gly 512 and Tyr 513. Sulfate-2 is bound near the upper edge of the interdomain cleft towards the N-terminal domain and interacts with Arg 20, Lys 29 and Lys 275. Arg 20 also interacts with the sulfate-3 that is positioned at the dimer interface and interacts with residues Arg 427 and Arg 428 from the helix $\alpha 16$ of the other monomer. The fourth sulfate ion is bound closer to the tetrameric interface and interacts with residues Gln 582, Asn 585 and Arg 589 from the regulatory helix ($\alpha 22$).

As mentioned earlier, the triple mutant crystallizes in two different crystal forms and the major difference between the two native apo-enzyme structures is in the position of sulfate-4. In the native2 structure, subunits C and D retain the sulfate-4 while in the A and B subunits sulfate-4 is absent and is apparently replaced by the binding of a single sulfate between the two regulatory helices. Residues Arg 589 and Arg 592 are contributed from both the A and B monomers to hold the sulfate at the molecular 2-fold axes that runs between and perpendicular to the regulatory helices. The molecular consequence of this

sulfate binding and apparent displacement of sulfate-4 from the A and B subunits, is that the relative positioning of these two monomers within the tetramer changes as they are drawn toward one another to better interact with the sulfate ion. In the native1 structure, the distance between the main-chain atoms of the regulatory helices across the tetrameric interface is approximately 8.0 Å. Upon binding of the sulfate ion between these helices, their separation decreases to 5.1 Å and the N-terminal Rossmann domain of one of the subunits (chain B) rotates by an angle of 1.5° towards the intra domain cleft.

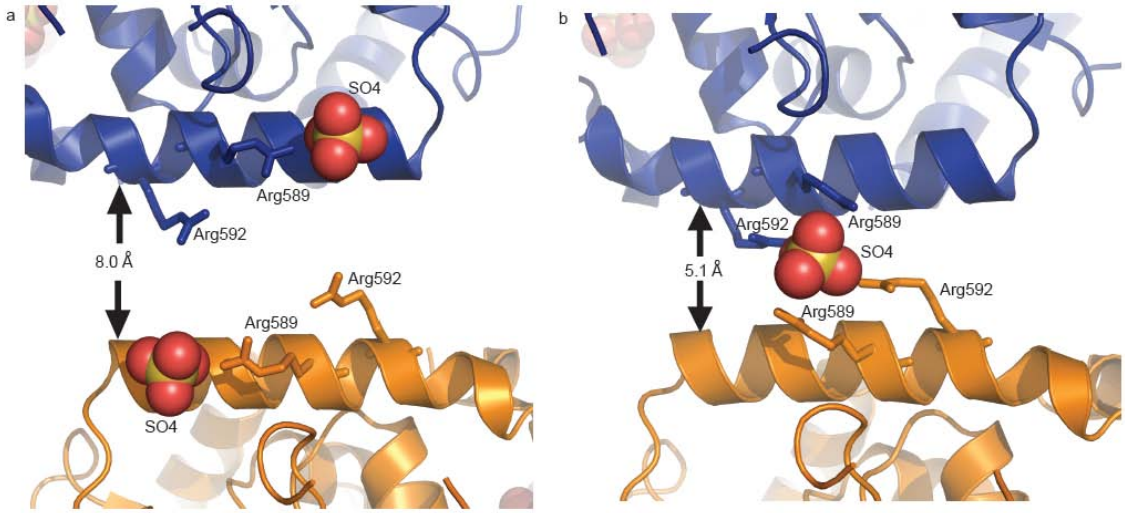

Figure 366. Sulfate binding at tetramer interface.
Comparison of sulfate binding at dimer-dimer interface of native-1(a) and native-2(b) structures. Ribbon diagram of the protein highlighting the interface with the bound sulfate represented in space filling models. The residues involved in binding are labeled.

2. Effect of sulfate on Gsy2p activity

To determine the effect of sulfate on the enzyme activity, we performed a sulfate titration of the yeast Gsy2p wildtype and arginine mutants under the standard assay conditions in the presence of 0-1M Li_2SO_4 (Figure 37).

The wild type enzyme shows an increase in the activity with Li_2SO_4 up to about 10 mM, whereupon further increase in the sulfate concentration has an inhibitory effect on the enzyme activity. Both the triple arginine mutant forms of yeast Gsy2p are also inhibited by sulfate, but with different concentration dependencies and are not inhibited to the same extent as the wild-type enzyme. The inflection points of the curve for the wild type and the R580A/R581A/R583A mutant occur at 200 mM sulfate, which is equivalent to the Li_2SO_4 concentration present in our crystallization conditions. It is worth noting that the inhibition curve is displaced to the right for the R587A/R589A/R592A mutant that lacks the two arginine residues involved in the binding of sulfate between the two regulatory helices. In the presence of saturating concentrations of glucose-6-phosphate, the arginine mutants show the same effects as that observed in the absence of glucose-6-phosphate. However with the wild type enzyme the initial increase in activity is not observed and there is a rightward displacement of the inhibition curve.

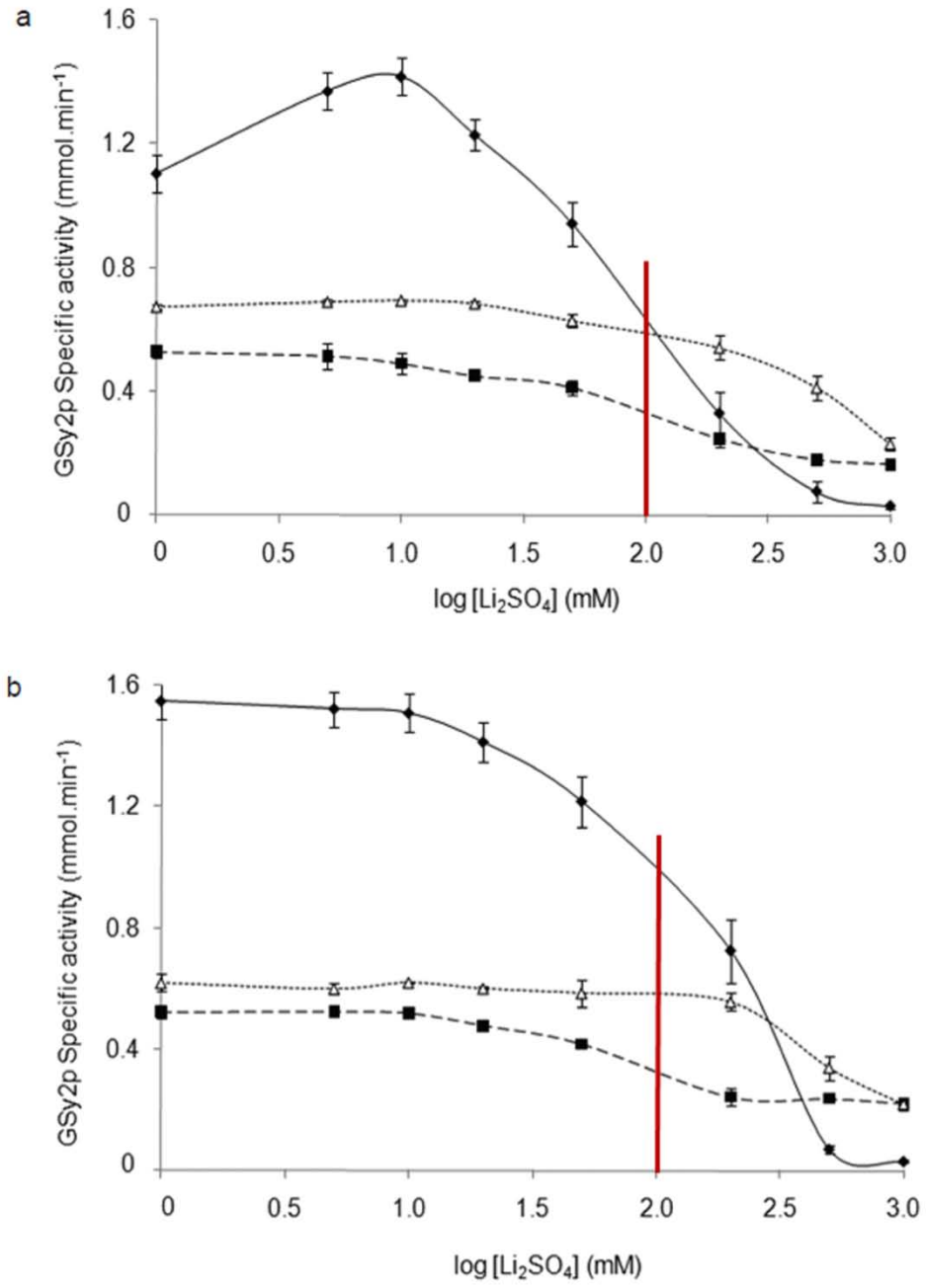

Figure 377. Effect of sulfate on Gsy2p activity.
The specific activity of wildtype (diamond), R580A/R581A/R583A (square) and R586/588/592A3 (open triangle) enzymes were measured in the absence (a) and presence (b) of saturating 7.2 mM glucose-6-phosphate and varying concentrations of Li_2SO_4 between 0-1M. The activity measurements were done with 4.4 mM UDP-glucose and 6.7 mg/ml glycogen. The concentration of sulfate used for crystallization is marked shown as red lines. The error bars represent the standard error of mean from three individual experiments.

DISCUSSION

A. Overall structure and oligomeric state

The basic folded architecture of each subunit in Gsy2p and the GT5 family members is conserved and includes two Rossmann domains (Figure 35). Within the N-domain, the GT3 and GT5 members have two major sequence insertions; the position of the insertions in relation to the structural elements of the core Rossmann domain is the same but both the lengths and relative three-dimensional positions vary. Between the β2 and β5 strands, the bacterial enzymes have a 35 amino acid insertion whereas in the yeast enzyme the insert is longer at about 60 amino acids. The second insert between the β7 strand and α10 helix has three helical segments that are positioned between the two Rossmann-domains, interacting closely with the C-terminal domain in the yeast Gsy2p. In the bacterial enzymes, this same insert is comprised of a two small helical segments located on the surface of the N-terminal domain.

The C-terminal domain of the yeast enzyme also has two unique sequence insertions and these insertions are present in all the mammalian enzymes. A 100 amino acid insertion is present between the second helix and second-strand of the canonical Rossmann-fold. The sequence identity of this insert across the different eukaryotic species is between 50-60%. Amongst the mammalian enzymes, the sequence identity in this region is even higher, at about 90%. A second small insert is present near residue 480 of Gsy2p and is also conserved across the eukaryotic species. The long inserts within the N- and

C- domains play essential roles for subunit oligomerization whereas the small C-domain insert is involved in nucleotide sugar donor binding.

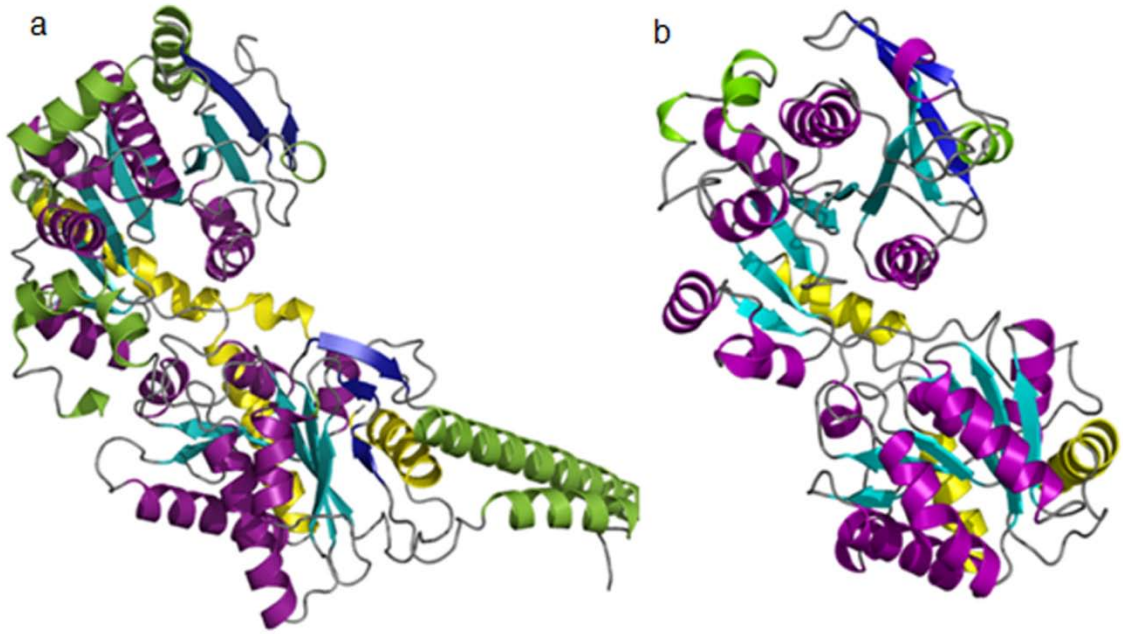

Figure 38. Comparison of GT3 and GT5 monomer.
Ribbon representation of the structure of GS monomer from yeast (a) and *E.coli* (b). The conserved secondary structural elements of the core Rossmann domains are represented in purple and cyan, the linker region in yellow and the unique eukaryotic inserts in blue and green.

B. UDP binding

In Gsy2p, the C-domain insertion that begins near residue 480 in Gsy2p creates an extended loop that redirects the main chain away from the active site region, shortens the following helix ($\alpha 18$) and provides sufficient space to facilitate the binding of the uridine ring within and beneath this loop structure (Figures 31 and 36). Repositioning of the helix and the preceding loop enables

the side chains of Phe 480 and Tyr 491 to form aromatic stacking interactions with the uridine ring.

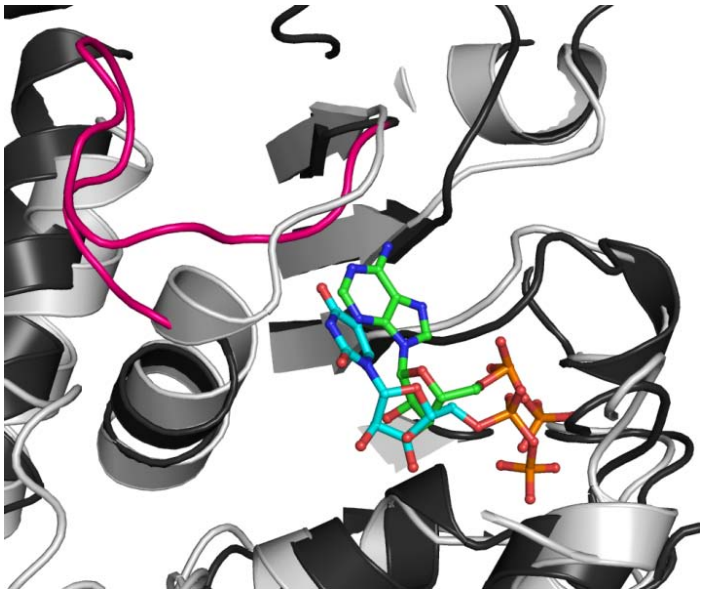

Figure 39. Nucleotide binding site in GS.
Ribbon representation of the superposed Gsy2p (dark grey) and Agrobacterium GS (light grey) structures. Gsy2p insert is highlighted in dark pink. The bound UDP and ADP are represented in ball and stick models in cyan and green respectively.

These aromatic side chains are conserved amongst eukaryotic GS enzymes. Remarkably, it appears that selectivity for the uridine base is achieved by the main chain of Lys 481, which is positioned to donate a single hydrogen bond to the exocyclic carbonyl group of the uridine ring (Figure 33). A similar interaction is not possible with CDP-sugars nor is the available space sufficient for any of the purine-based nucleotide sugars. The pyrophosphate group of the UDP interacts with the side chains of Arg 320 and Lys 326, residues that are

conserved in all the glycogen synthase enzymes. In addition to these side chains, the pyrophosphates of UDP also interact with the helical dipole provided by helix $\alpha 19$ and participate in hydrogen bonding interactions with the side chain of Thr 514. In the apo and maltodextran bound R580A/R581A/R583A structures, a sulfate ion is bound at the position of the UDP β-phosphate group.

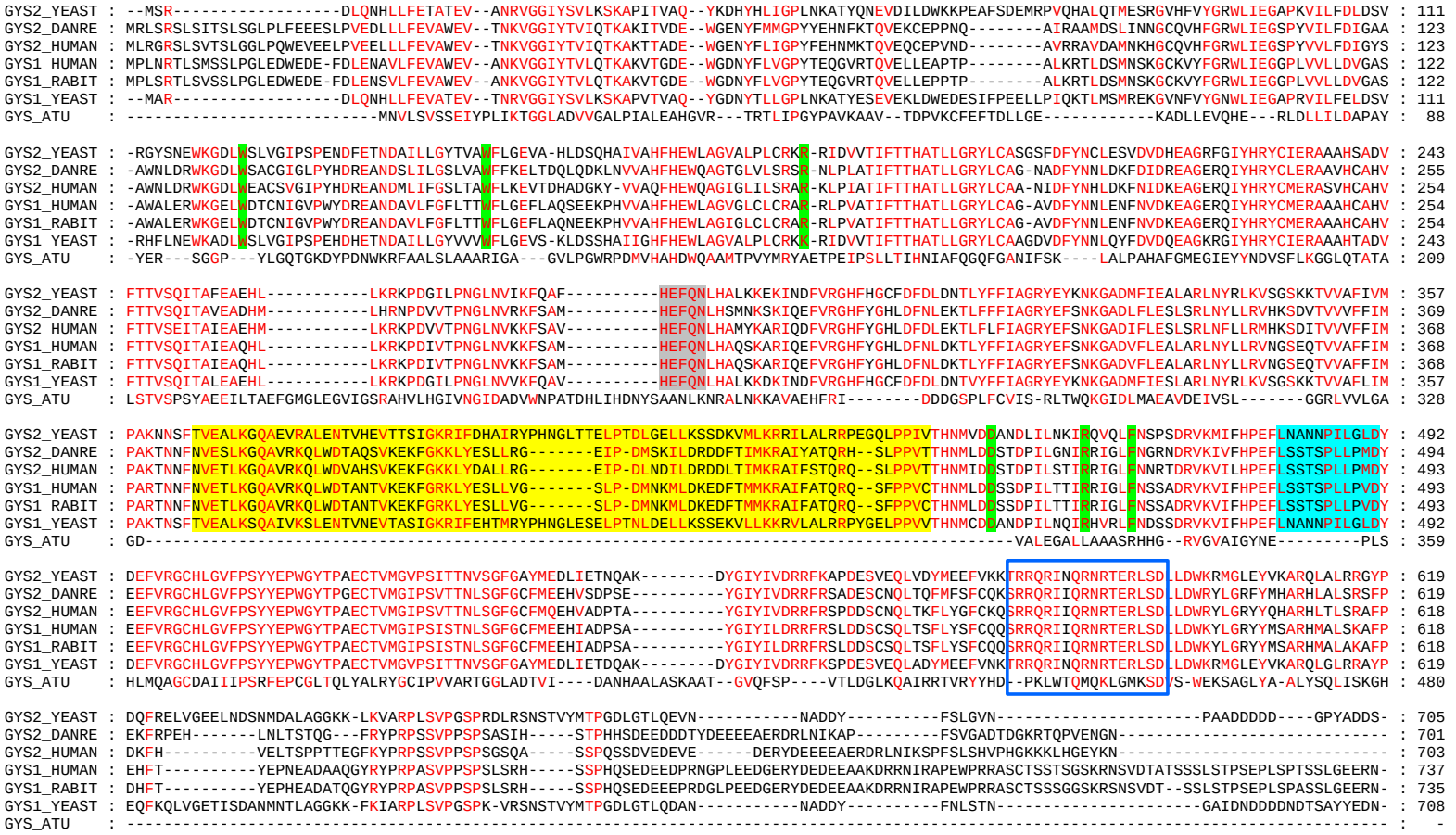

Figure 40. Sequence alignment of GS.
Residues represented in red show >90% conservation across various species. The C-domain long helix and 480 loop insertions are highlighted in yellow and cyan. Residues involved in maltodextran binding, stabilizing the activated conformation are highlighted in green and grey. The regulatory helix is marked by the blue rectangle.

C. Catalytic mechanism

GS enzymes are metal ion independent retaining type glycosyl transferases whose mechanism of action remains elusive. One study with the human muscle GS enzyme suggested a S_N2 type reaction with the two glutamate residues of the conserved E-X_7-E motif acting as the nucleophile and general acid/base[38]. In our structures, the proposed nucleophile, Glu 509 of the Gsy2p is turned away from the active site and interacts with the side chain of Lys 326, which might be a preferred mode of interaction in the absence of a donor sugar attached to the nucleotide. Rather than acting as a nucleophile, Glu 509 might be involved in stabilizing the pyrophosphate of the leaving group by correctly positioning the charged residues Arg 320 and Lys 326. Alternatively, a different residue, such as Tyr 513, might function as the nucleophile in the eukaryotic organisms. The second glutamate residue in the E-X_7-E motif - Glu 517 - is conserved only in the eukaryotic and archaeal enzymes that are capable of using UDP-glucose as the substrate. In our UDP bound structure, Glu 517 is positioned to interact with the 2' and 3' hydroxyls of the uridine ribose moiety and is, therefore, unlikely to act as the general acid/base for the catalytic reaction. However, an unusual cluster of polar residues comprised of His 166, His 168, Thr 191, His193 and Thr 245 form a hydrophilic core within the N-terminal domain and appear to link the active site region with solvent. These residues are conserved across all the GS enzymes. We speculate that His193 might be the general base and that the remaining residues could serve as a proton wire

shuttling protons from the active site to solvent without the need to reset the protonation state of the active site through release and rebinding of the growing acceptor molecule. A possible role for high efficiency proton shuttling would be to provide a mechanism through which the processivity of eukaryotic GS enzymes is enhanced. Consistent with this proposal, mutation of the central residue in this hydrogen bonding network, His 166, reduces catalytic activity to 1% of wild-type[62].

D. Maltodextran binding sites

Binding of maltodextran to the same two sites in subunits A and B of the R580A/R581A/R583A mutant was observed both in soaking and co-crystallization experiments with maltooctaose. The two sites are found on the surface of the N-terminal (site 1) and C-terminal (site 2) Rossmann-fold domains (Figure 37c). The hydrophobic and the polar residues that constitute these maltodextran binding sites are highly conserved across the eukaryotic species (Figure 40). It is interesting to note that maltodextran binding sites equivalent to our site 1 position are also found in glycogen phosphorylase (termed the glycogen storage site) and in the *E.coli* glycogen synthase enzymes (Figure 37). In each case, the locations are approximately similar, but the maltodextran binding sites share no sequence or structural homology and are formed from unique elements of secondary structure. The site 1 region on the C and D subunits are utilized in crystal packing contacts and are unavailable to interact with maltodextran in the low activity state structure.

The second maltodextran binding site is located in the C-terminal domain and appears to be unique to the eukaryotic enzymes. Like the site 1 binding cleft, this site is also lined with conserved residues and is located at the structural boundary between the long helices that form the oligomeric interface and the C-terminal Rossmann domain. Interestingly, three missense mutations reported in the human liver GS enzyme that lead to glycogen storage disease type-0 occur amongst conserved residues that are located in close proximity to this second maltodextran binding site.

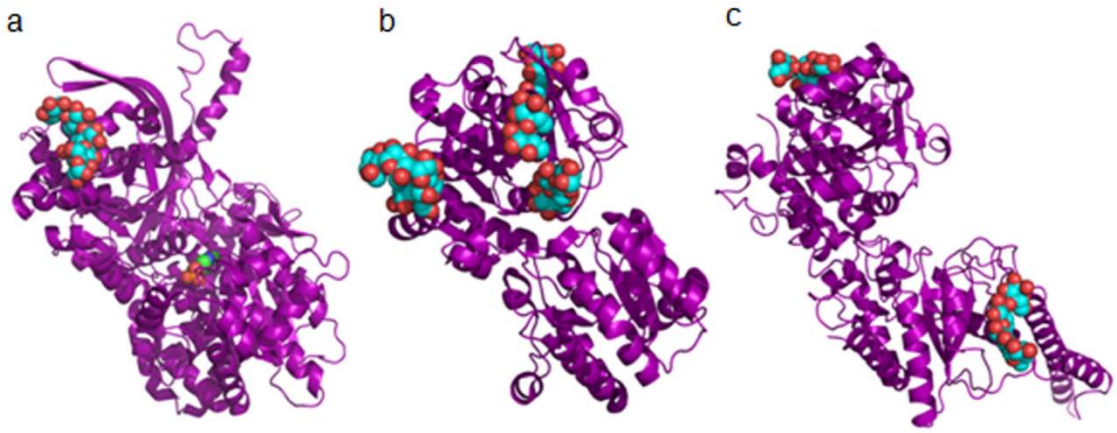

Figure 40. Maltodextran binding sites in GTB enzymes.
Ribbon representation of themonomer structure of rabbit muscle glycogen phosphrylase (a), *E.coli* GS (b) and yeast Gsy2p (c). The bound maltodextran polymers are represenced in space fill models.

In Gsy2p, the 4' ends of the polymers bound at each site are between 30 and 35 Å from the catalytic site (Figure 41c) and as such do not appear to represent sites appropriate for proper positioning of the acceptor end of the growing glycogen chain. Instead, consistent with the multiple polymer binding site model for rabbit GYS1[44], we propose that these are non-catalytic, docking sites

for glycogen on the synthase enzyme, which either keeps the enzyme in close proximity to the polymer for efficient catalysis or provides a convenient docking site while the branching enzyme acts on chains that have reached lengths appropriate for introduction of a branch point. The multiple binding sites may explain the propensity of glycogen synthase to remain tightly associated with the glycogen particle and may suggest a role for increasing the processivity of glycogen synthase. The maltodextran binding site mutants support this model as the mutation of critical residues in these sites decreases the V_{max} and the $S_{0.5}$ glycogen of the enzyme. In the presence of glucose-6-phosphate, the $S_{0.5}$ glycogen is decreased by half for the wild type enzyme whereas it is increased two to three times for the mutants (Table 7). Since there are no changes in the catalytic and allosteric sites of the wild type and mutant Gsy2p, the allosteric effects of glucose-6-phosphate should be similar. The higher activity of the enzyme in the presence of glucose-6-phosphate demands the need for greater amount of the polymer substrate for extension. Since the wild type enzyme is tightly associated with glycogen, there is always excess glycogen available. However, the mutants are not tightly bound with glycogen and hence we need an increased local concentration of the substrate to fulfill the demands of the activated enzyme, which is reflected in the increase in the $S_{0.5}$ glycogen.

D. Activation by glucose-6-phosphate

The 100 amino acid C-domain insertion of Gsy2p fold as an extended coiled-coil, traverses the dimer-dimer interface and forms two four-helical

bundles comprising the majority of interactions in the oligomer assembly. The helix α-16 of the coiled-coil domain interacts with helix $\alpha 2$ and strand $\beta 3$ within each N-domain insertion in the low activity state - R580A/R581A/R583A structures (Figures 25 and 30a). This interaction between the N-terminal and C-terminal insertions represents the primary points of contact between the large oligomerization insertion and the remainder of the enzyme in these structures. This assembly positions the regulatory helices far away from each other in the dimer (~60 Å). However, when the two dimers are assembled back-to-back in the tetramer, these same regulatory helices are now positioned opposite their partner helices at a distance of approximately 8 Å in the R580A/R581A/R583A structures (Figure 26). The interactions of α-16 helix with the N-terminal insertion are replaced by interactions with residues 482-487 when the helical domains collapse toward the central interface in response to glucose-6-phosphate binding (Figure 30b).

As might be expected for an allosteric molecule that induces such large conformational changes, the interactions between the enzyme and glucose-6-phosphate are extensive and involve residues from more than one subunit (Figure 28). In particular, the 6-phosphate is completely enclosed in a binding pocket comprised of five residues (His 286, Lys 290, His 500, Arg 583 and Arg 587) that are conserved across eukaryotes. The C1’ and C2’ hydroxyl groups of the glucose moiety forms hydrogen bonds to Gln 283 and His 280, respectively. Unlike the majority of the phosphate-binding site, the glucose binding site is not well-formed prior to binding. In particular, residues 278 to 284 are disordered in

the basal state and adopt a stable conformation only when the activator is bound (Figure 29). Consequently, most of the conformational changes are brought about by the interactions surrounding the glucose moiety, including those contributed by the opposing subunit. The phosphoryl group serves as a strong anchor upon which to build the necessary interactions to drive the conformational changes.

Based on the changes in local structure, it would appear that the trigger for these conformational changes are the interactions contributed by His 280, Gln 283 and Asn 284, the latter of which hydrogen-bonds with itself across the subunit interface. Although residues His 286 and Lys 290 also contribute hydrogen bonds to the bound glucose-6-phosphate, their relative positions change very little upon glucose-6-phosphate binding. Rather the ordering of the loop between 278 and 284 and drawing of this loop into interactions across the regulatory interface stabilizes the activated conformation by pushing the helices apart and inducing rigid body rotations and translations in order to achieve optimal interactions for His 280, Gln 283 and Asn 284. The only other residue that may assist in stabilizing this new interface is the side-chain of Arg 581, which forms weak and variable interactions with the main-chain carbonyl oxygen atoms of residues 276 and 277 in the opposing subunit. However, given its comparatively weak electron density and the mutagenesis data on Arg 581 (Table 6), this residue is likely more important in determining the structural response to phosphorylation than in contributing to the stability of the activated conformation.

The reorganization of the subunit interface likely leads to increased enzymatic activity because the active site clefts are now freely open for substrate access and the N-terminal Rossmann domains at the periphery of the tetramer are no longer anchored by their interactions with helix α16 (Figure 40). Structural studies with glycogen synthase from *E.coli* have demonstrated that the transition between the resting open and catalytically-poised closed conformation of the enzyme involves a 15° inter-domain rotation[72]. The relative domain orientation of Gsy2p in the low activity state is most similar to the open-form of the *E.coli* enzyme and domain closure similar to the *E.coli* enzyme is hindered by the interactions between the N-terminal domains and helix α16. However, upon activator binding, the resulting changes in subunit positioning frees the N-terminal domains from the constraining interactions with α16 and the enzyme can adopt the opened and closed domain conformations required for catalysis (Figures 40 and 41). Consistent with this hypothesis, the N-terminal domains of the A, C and D subunits are actually rotated an additional 4.6° away from the interface in the activated structure than in the R580A/R581A/R583A structure (Figure 42). Interestingly, the B subunit is 12.8° more closed than in the basal state structure. This domain closure appears to be precipitated by serendipitous binding of an additional glucose-6-phosphate molecule within the active site that triggers the domain closure. No such variation in domain orientation was observed in the low activity structure, which is consistent with our contention that the subunit positioning in the activated state facilitates catalytically important conformational changes.

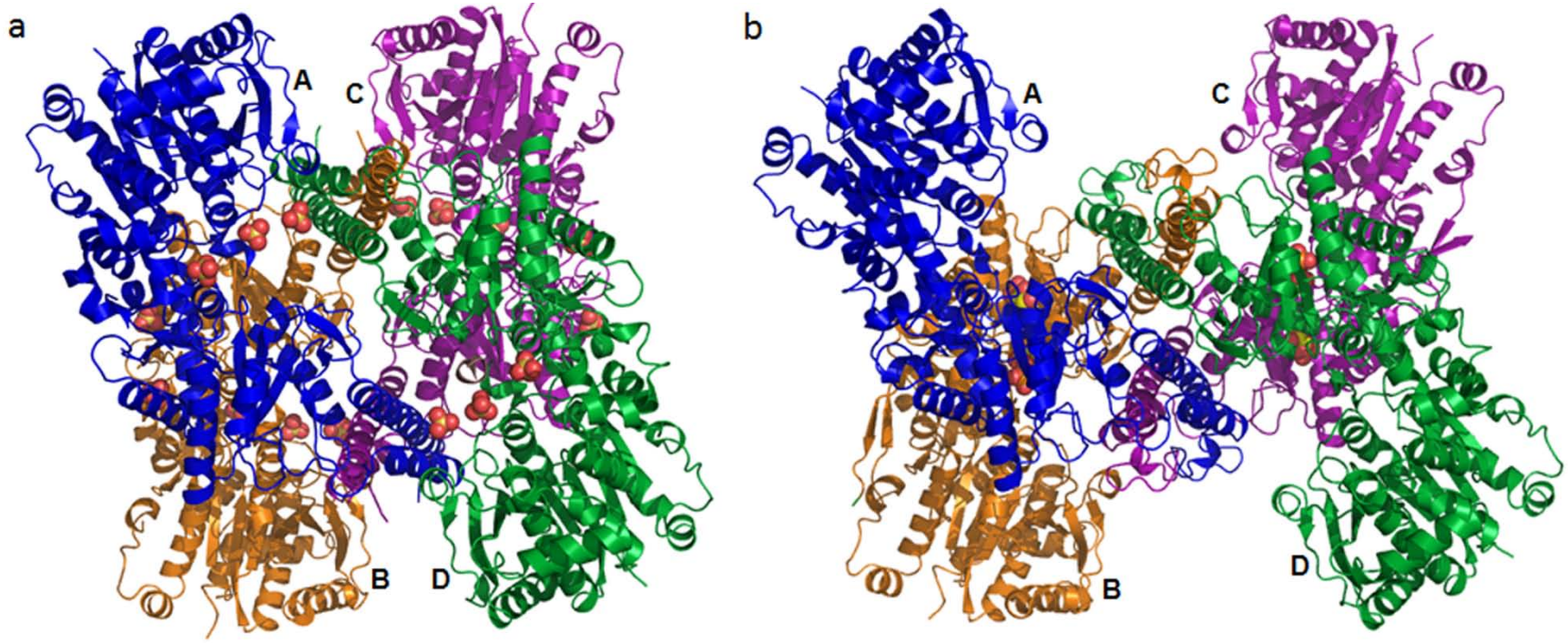

Figure 41. Comparison of Gsy2p conformations – Active site.
(a) Ribbon representation of Gsy2p low activity state(R580A/R581A/R583A tetramer) structure with the bound sulfates in space filling models. (b) Ribbon representation of Gsy2p activated conformation (R589A/R592A tetramer) with the bound glucose-6-phosphate in space filling models Different colors are used to represent different subunits.

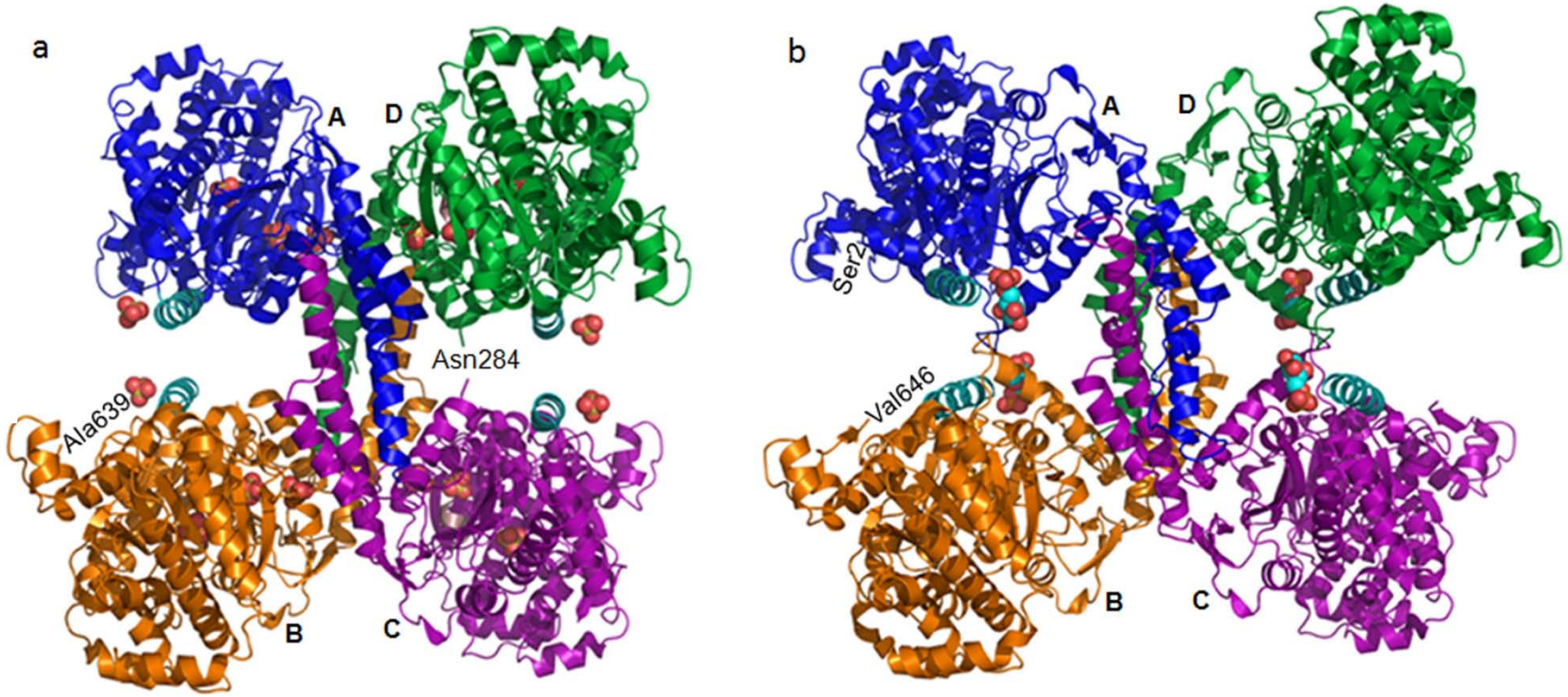

Figure 42. Comparison of Gsy2p conformations - Regulatory helix interface.
(a) Ribbon representation of Gsy2p low activity state(R580A/R581A/R583A tetramer) structure with the bound sulfates in space filling models. (b) Ribbon representation of Gsy2p activated conformation (R589A/R592A tetramer) with the bound glucose-6-phosphate in space filling models Different colors are used to represent different subunits.

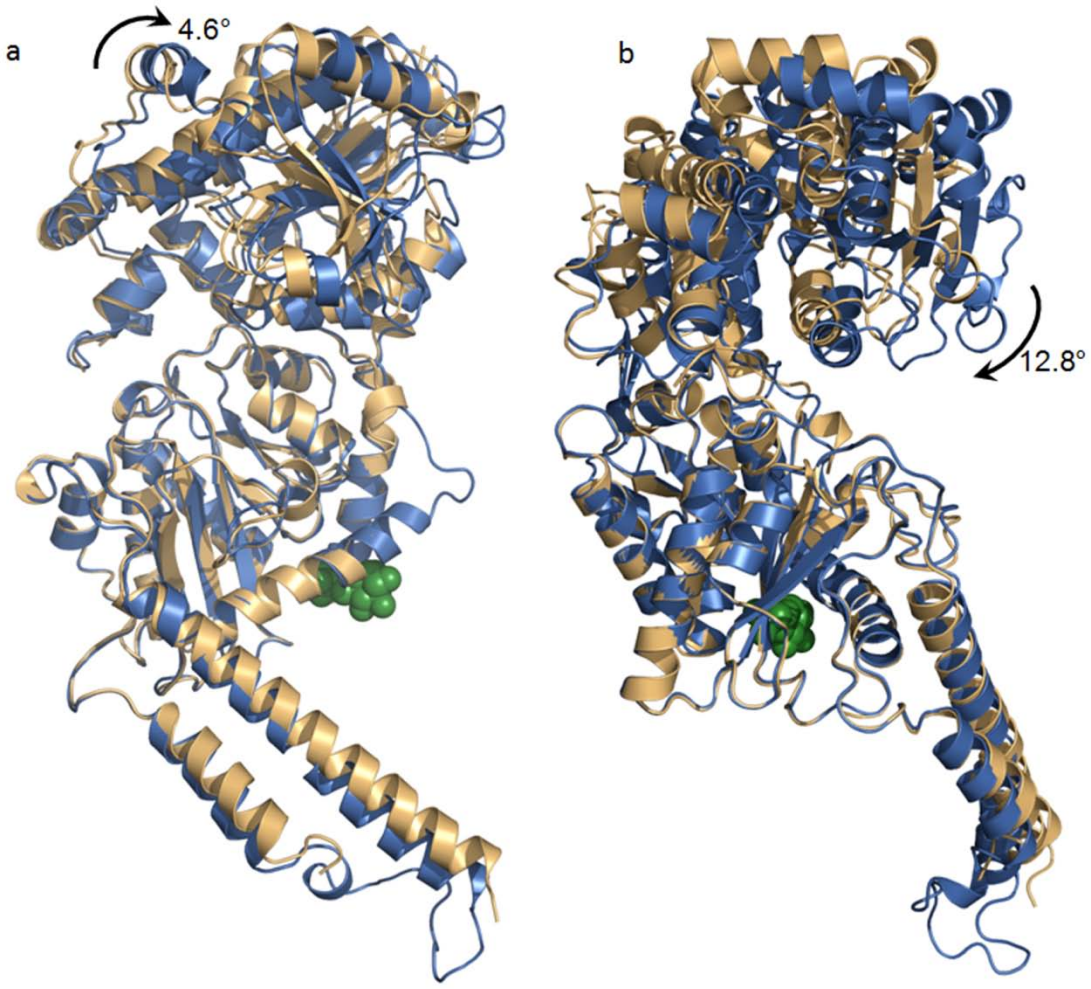

Figure 43. Comparison of the R580A/R581A/R583A and R589A/R592A activated state conformations.
Ribbon representation of the superposed basal state and activated state subunit A (a) and subunit B (b) with the bound glucose-6-phosphate in space filling model. Different colors represent different conformations: orange-low activity, blue-high activity. The alignment was generated by superposing the C-terminal Rossmann domain (residues 312-357 and 453-577) using the superpose program of the CCP4 suite.

E. Role of Regulatory helix arginines in conformational transition

The mutagenesis and intein-mediated phospho-peptide ligation data suggest the arginine residues within the regulatory helix play specific roles in the conformational selection process that ultimately control the different activity states. First, it is clear from the double mutants and the glucose-6-phosphate activated structure that Arg 583 and Arg 587 are necessary and sufficient for conferring glucose-6-phosphate sensitivity. But rather than directly triggering the conformational changes upon activator binding, their roles are to provide anchoring interactions for the phosphate so that the interactions conferred by residues 280-284 with the glucose moiety can trigger the conformational changes. Second, in the absence of glucose-6-phosphate, charge neutralization of the other four arginines is in some way responsible for shifting the equilibrium from an activated state toward a less active state.

The mutagenesis data show that neither Arg 580 nor Arg 581 make significant contributions toward establishing the basal activity state or toward stabilizing the glucose-6-phosphate activated state (Table 6). However, prior work[62] and the current intein-mediated phospho-peptide data demonstrate that their contributions are critical to stabilize the inhibited state following C-terminal phosphorylation. Therefore, these two arginine residues likely contribute to the binding of the phosphorylated Thr 668 residue and/or help hold the enzyme in a less active state. On the other hand, Arg 589 and Arg 592 seem critical to

establishing the activity of the basal state, since their mutation to alanine reduces the basal activity of the enzyme to levels similar to the phosphorylated state.

F. Inhibition by C-terminal phosphorylation

The kinetic studies demonstrate that sulfate concentrations similar to that utilized in the crystallization of the R580A/R581A/R583A mutant has an inhibitory effect on the enzyme activity. The kinetic effects of sulfate are not simply the result of non-specific inhibition by high ionic strength since neither of the two triple mutants exhibit the moderate activation or full inhibition shown by the wild-type enzyme in response to varied sulfate concentrations. Thus, the behavior of the triple mutants toward sulfate is similar to that toward the physiological activators and inhibitors; their response is muted when compared to the wild-type enzyme. Since sulfates are known phosphomimics, we suggest that the inhibitory phosphates bind at or near the regulatory helices, similar to the sulfate binding seen in the low activity state structure and result in charge neutralization of these arginine residues. Though the precise location of the phosphate at the regulatory helix is not known, we speculate that phosphate binding would result in constraining the positions of the regulatory helices and limit the conformations accessible to the inhibited state. It is possible that the low specific activity, the increase in the K_M for UDP-glucose and the changes in $M_{0.5}$ for glucose-6-phosphate exhibited by the phosphorylated state of Gsy2p are indirect effects of these conformational restraints.

Either an intra- or intersubunit interaction with the phosphothreonine residue at position 668 would fulfill the requirement for charge neutralization of the regulatory arginines. However, an intersubunit interaction would add an additional conformational constraint in the form of a locking strap running from the N-terminal domain of the opposing subunit, where the electron density of the low activity structure terminates at residue 640 approximately 25-30 Å from the regulatory interface, across the interface to the regulatory helix located on the opposing C-terminal domain (Figure 42). Indeed, the additional residues observed at the C-terminus of the glucose-6-phosphate activated structure of the R589A/R592A mutant are directed precisely toward the regulatory interface (Figure 42). The basal activity of the R589A/R592A double mutant support an intermolecular locking mechanism as its lower activity ratio does not approach the activity ratio of the phosphorylated wild-type enzyme and an intermolecular mechanism is also supported by the stronger inhibition of the longer phosphopeptide construct (Table 6). Binding of glucose-6-phosphate presumably disrupts these locking interactions through the conformational changes across the regulatory interface and throughout the tetramer.

While phosphorylation of the Thr668 residue alone decreases the activity ratio of Gsy2p in a manner comparable to that of the in vitro phosphorylated enzyme, this is not sufficient to completely block the access of the glucose-6-phosphate to the activity state. The $M_{0.5}$ of glucose-6-phosphate for the in-vitro phosphorylated form of Gsy2p is 800 µM, which is 20 times greater than the value for the dephosphorylated form[62]. However, ligation of either the 35-mer or

the 49-mer Thr668 phosphopeptide to the wild type Gsy2p core only increased the $M_{0.5}$ by three-fold . This suggests that additional phosphates at the serine residues are necessary for blocking the access of the activator to the allosteric site and one possibility is that one of the inhibitory phosphates could interact with the 6-phosphate binding site of the activator.

G. Regulatory model for Gsy2p

Based on the current structural and biochemical work on Gsy2p, we provide molecular insight into the nature of the three activity states of Gsy2p, which span a 20-fold range of specific activity and the apparent driving forces that shift the conformational equilibrium between these states (Figure 44). Rather than three distinct states, we believe that there is a dynamic equilibrium between the three states and the specific activity is an indication of the conformational flexibility available for a particular enzyme to exist in each state. This raises the probability of some intermediate conformation and indeed the R580A/R581A/R583A structures are examples of such intermediate conformation between the basal activity intermediate state and the low activity inhibited state.

In the highest activity (R) state, binding of glucose-6-phosphate adjacent to the regulatory helices pushes the helices apart and results in domain rotations and translations that frees the conformational restraints, opens up the active site and facilitates interdomain motions that are necessary for catalysis. The

activated conformation is stabilized by the ordering of residues 277-284 that not only interact with the activator molecule within its own subunit but also with the associated monomer.

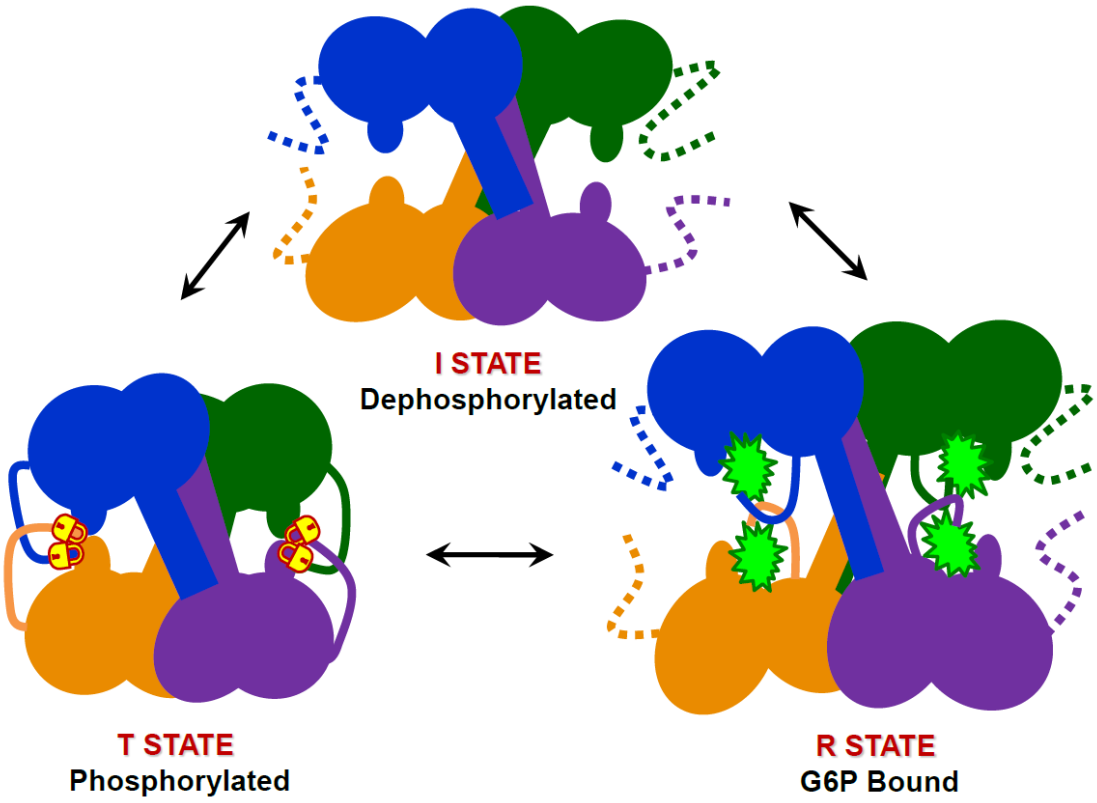

Figure 44. Three state structural model for Gsy2p.
Three state structural model for regulation of Gsy2p. Different colors are used to represent the individual subunits of the tetramer and the dotted lines represent the disordered C-terminal region in the intermediate and activated states.

Though we lack solid structural data on the inhibited (T) state conformation, based on the kinetic data and the structural information from the triple mutant, we propose that binding of phosphate at or near the regulatory helix would result in structural collapse of the helices, pull the subunits closer and

restrain the domain motion involved in catalysis. While a single phosphate group at the Thr 668 position is sufficient to inhibit the enzyme activity, we suggest that the additional phosphates and the conserved hydrophobic patch in the C-terminal region provide a locking mechanism to stabilize the T state conformation. Based on the regulatory helix movement seen in the R580A/R581A/R583A structures, we would expect that the dephosphorylated wild-type enzyme to have an intermediate level of flexibility.

Since the secondary structural elements that are involved in oligomerization and activity regulation are highly conserved between the mammalian GS enzymes and yeast Gsy2p, this work establishes the paradigm for regulatory control of the mammalian glycogen synthases. The major difference between the mammalian and the yeast enzymes is the presence of additional phosphorylation sites in the C and N-terminal regions. As mentioned earlier, the additional phosphorylation sites could provide further locking interactions to stabilize the inhibited conformation. In all our structures, the N-terminus is positioned similar to the C-termini at a distance of 40 Å from the regulatory helices, but the spatial positioning in the tetramer is such that any N-terminal extensions would have easy access to the regulatory helices (Figure 41). Hence it is possible that the N-terminal phosphorylation sites also bind to the regulatory helices and restrain the conformations as we predict for the C-terminal sites. While the specific activity of the different conformational states of yeast Gsy2p vary by approximately 20-fold[62], the different activity states of the mammalian enzymes show a much higher variation in the specific activity (> 200

fold)[95,96]. It is possible that the complex nature of the mammalian tissues demand the presence of additional phosphorylation sites to transmit and integrate signals from different inputs for manipulating and fine tuning the GS enzyme activity over a large range to meet the physiological requirement of the specific involved tissues.

CONCLUSIONS

The structure of yeast Gsy2p in the low activity and glucose-6-phosphate bound activated conformations provides the molecular and structural basis of regulation of eukaryotic GS enzymes. The most intriguing feature of Gsy2p is the tetrameric arrangement mediated by the helical insert in the C-terminal domain; this unique arrangement facilitates extensive conformational flexibility and effective communication between the active site of the individual subunits, which are essential for the regulation of enzymatic activity. The conserved arginines are located in a single helix and are positioned at a molecular two-fold of the tetramer assembly. The arginines function as sensors for both the activator and the inhibitory signals and are involved in fine tuning the enzyme activity in response to the physiological demands of the cells or tissues. Binding of glucose-6-phosphate at the tetrameric interface adjacent to the regulatory helix relieves the conformational restraints imparted on the subunits and thus facilitates the intra domain rotations necessary for enzyme activity. Inhibitory phosphate binding at or near the regulatory helix would lock the enzyme in a restrained conformation and thus impede these movements. A small insert if the C-terminal domain at the substrate binding site is involved in mediating the UDP-glucose specificity of the eukaryotic enzymes. Further, the presence of sugar polymer acceptor binding sites on the enzymes surface help the tight association of the enzyme with the substrate and increases the processivity of the enzyme.

FUTURE DIRECTIONS

We have collected a 2.5 Å dataset of the UDP and maltodextran bound complexes of the activated conformation of Gsy2p and expect that these structures would provide additional clues on the mechanism of enzyme activation induced by the rotational movements of the domains. The structures of the complexes were solved by molecular replacement using the R589A/R592A glucose-6-phosphate bound tetramer as the search model. Preliminary analysis of the structures showed no significant changes in the UDP binding, but we do observe additional maltodextran polymer binding sites in the activated conformations. We have also collected a 2.5 Å dataset of the activated conformation of the R580A/R581A mutant in the I222 space group and solved the structure using the R589A/R592A tetramer as the search model. Analysis of our current model did not show any significant changes in the activated conformation.

Our next structural work will focus on both the state I and state R. We have some crystals hits for the apo form of R589A/R592A mutant; optimization of the hit and solution of the apo structure would give substantial information on the inhibited conformation. A new screen for the apo form of the wild type, R580A/R581A and the T667 phosphopeptide ligated wild type has to be performed to gain additional information on the intermediate and activated conformations. Site-directed mutageneis and enzyme kinetic studies have to be done to confirm the role of the conserved polar His cluster in catalysis. The

glucose-6-phosphate bound in the active site of the B subunit in the activated conformation provides some insights on the catalytic nucleophiles and these have to be tested by mutagenesis and enzyme kinetic approaches.

The ultimate goal of this project is to extend the studies to the mechanism of regulation of the mammalian GS enzymes. Since the yeast Gsy2p shares 50% sequence identity to the mammalian species and the secondary structural elements are conserved between the two species, we would expect to solve the mammalian GS structures by molecular replacement using the yeast structure as the search model.

APPENDICES

1. SOLVE/RESOLVE script for determining heavy atom position and density modification

```
setenv   SYMINFO /usr/local/lib/solve/syminfo/lib
setenv   CCP4_OPEN UNKNOWN
setenv   SOLVETMPDIR /var/tmp
setenv   SYMOP /usr/local/lib/solve/symop.lib
setenv   SYMINFO /usr/local/lib/solve/syminfo.lib
unlimit
#output and log file  solve an MIR problem
solve<<EOD  >  solve.log
logfile   solve.logfile
#resolution and symmetry
resolution   20  5.0
cell  96.5   166.3   121.1   90.0   103.4   90.0
symfile    /usr/local/lib/solve/p21.sym
#native data
readdenzo
premerged
read_intensities
fixscattfactors
rawnativefile   native1.sca
#derivative 1
derivative 1
label  deriv  1 TA36
atom  ta
rawderivfile   tbr36.sca
nsolsite_deriv   6
```

```
#derivative 2
derivative 2
label  deriv  2 TA44
atom  ta
rawderivfile   tbr44.sca
nsolsite_deriv   5
SCALE_NATIVE
SCALE_MIR
ANALYZE_MIR
SOLVE
EOD
#resolve for density modification
resolve  << EOD  >  resolve.log
solvent_content   0.58
nobuild
EOD
```

2. PHENIX script for partial model phasing

```
phenix.autobuild Maps_only=True \
model=partial_model2.pdb \
phenix.simple_ncs_from_pdb partial_model.pdb \
data=solve.mtz \
input_hires_file=nativet1.mtz \
```

3. Pymol script for generating figures with difference electron density maps around bound ligand

```
set bg_rgb,[1.0, 1.0, 1.0]
set ray_trace_fog, 0
cmd.space('cmyk')
set two_sided_lighting,1
set cartoon_fancy_helices,1
set ray_shadows, 0
set stick_radius=0.2
set stick_ball=on
set stick_ball_ratio=1.2
set mesh_radius = 0.025
load 589a2_g6pa.pdb, g6p
cmd.dss("g6p")
load 589a2.pdb, gsy2
cmd.dss("gsy2")
set_color grey, (0.7,0.7,0.7)
select monA, (resi 2-639 and chain A)
select monB, (resi 2-639 and chain B)
hide all
show cartoon, gsy2
alter 280-301/, ss='H'
```

```
alter 599-611/, ss='H'
rebuild
select surres, (resi 587, resi 583, resi 580, resi 581, resi 500, resi 286, resi 290,
resi 283, resi 284, resi 280, resi 277, resi 281)
show stick, g6p
show stick, surres
hide ("surres" and name n+c+o)
util.cba (130,"monA")
util.cba(70,"g6p")
util.cba(30,"monB")
load 589a2_g6pa.ccp4, map1
isomesh diffmap, map1, 1.0, (g6p), 1, carve=2.1
set_color cyan=[0.1, 1.0, 1.0]
color cyan, diffmap
deselect
set_view (\
    -0.085544080,   -0.994874537,   -0.053918786,\
     0.831045389,   -0.101099811,    0.546939194,\
    -0.549586892,    0.001979203,    0.835435867,\
     0.000000000,    0.000000000,  -79.028060913,\
    38.875000000,  -26.892000198,   41.817001343,\
    71.476737976,   86.579383850,    0.000000000 )
```

4. Pymol script for generating figures with ligands and intreracting residues represented in space filling models and lines respectively

```
set bg_rgb,[1.0, 1.0, 1.0]
set ray_trace_fog, 0
cmd.space('cmyk')
set cartoon_smooth_loops, 1
set two_sided_lighting,1
set cartoon_fancy_helices,1
set ray_shadows, 1  # set 0 for interior shot
set antialias, 1
set line_width = 6.0
load nativet1.pdb, R580final
cmd.dss("R580final")
hide all
select Aso4, (resn SO4 and chain A)
select Bso4, (resn SO4 and chain B)
show spheres, Aso4
util.cba(3,"Aso4")
show spheres, Bso4
util.cba(3,"Bso4")
select monA, (resi 2-639 and chain A)
color blue, monA
```

```
show cartoon, monA
select monB, (resi 2-639 and chain B)
color orange, monB
show cartoon, monB
select argres, ( resi 589, resi 592)
show lines, argres
select monC, (resi 2-639 and chain C)
hide monC
select monD, (resi 2-639 and chain D)
hide monD
deselect
set_view (\
     0.261689752,    0.004401172,   -0.965144753,\
     0.593401134,   -0.789389372,    0.157295302,\
    -0.761181295,   -0.613876283,   -0.209186524,\
     0.000000000,    0.000000000, -113.639129639,\
    20.062999725,   48.997001648,   21.891000748,\
    85.909431458,  141.368804932,    0.000000000 )
```

REFERENCES

1 Melendez, R., Melendez-Hevia, E. & Canela, E. I. The fractal structure of glycogen: A clever solution to optimize cell metabolism. *Biophys J* **77**, 1327-1332, (1999).

2 Melendez-Hevia, E., Waddell, T. G. & Shelton, E. D. Optimization of molecular design in the evolution of metabolism: the glycogen molecule. *Biochem J* **295 (Pt 2)**, 477-483 (1993).

3 Kirkman, B. R. & Whelan, W. J. Glucosamine is a normal component of liver glycogen. *FEBS Lett* **194**, 6-11, (1986).

4 Fontana, J. D. The presence of phosphate in glycogen. *FEBS Lett* **109**, 85-92, (1980).

5 Tagliabracci, V. S. *et al.* Laforin is a glycogen phosphatase, deficiency of which leads to elevated phosphorylation of glycogen in vivo. *Proc Natl Acad Sci U S A* **104**, 19262-19266, (2007).

6 Krisman, C. R. & Barengo, R. A precursor of glycogen biosynthesis: alpha-1,4-glucan-protein. *Eur J Biochem* **52**, 117-123 (1975).

7 Alonso, M. D., Lomako, J., Lomako, W. M. & Whelan, W. J. Catalytic activities of glycogenin additional to autocatalytic self-glucosylation. *J Biol Chem* **270**, 15315-15319 (1995).

8 Rodriguez, I. R. & Whelan, W. J. A novel glycosyl-amino acid linkage: rabbit-muscle glycogen is covalently linked to a protein via tyrosine. *Biochem Biophys Res Commun* **132**, 829-836, (1985).

9 Smythe, C., Caudwell, F. B., Ferguson, M. & Cohen, P. Isolation and structural analysis of a peptide containing the novel tyrosyl-glucose linkage in glycogenin. *EMBO J* **7**, 2681-2686 (1988).

10 Caudwell, F. B. & Cohen, P. Purification and subunit structure of glycogen-branching enzyme from rabbit skeletal muscle. *Eur J Biochem* **109**, 391-394 (1980).

11 Bates, E. J., Heaton, G. M., Taylor, C., Kernohan, J. C. & Cohen, P. Debranching enzyme from rabbit skeletal muscle; evidence for the location of two active centres on a single polypeptide chain. *FEBS Lett* **58**, 181-185 (1975).

12 Roach, P. J. Glycogen and its metabolism. *Curr Mol Med* **2**, 101-120 (2002).

13 Greenberg, C. C., Jurczak, M. J., Danos, A. M. & Brady, M. J. Glycogen branches out: new perspectives on the role of glycogen metabolism in the integration of metabolic pathways. *Am J Physiol Endocrinol Metab* **291**, E1-8,(2006).

14 Ferrer, J. C. *et al.* Control of glycogen deposition. *FEBS Lett* **546**, 127-132, (2003).

15 Bouche, C., Serdy, S., Kahn, C. R. & Goldfine, A. B. The cellular fate of glucose and its relevance in type 2 diabetes. *Endocr Rev* **25**, 807-830, (2004).

16 Cheatham, B. GLUT4 and company: SNAREing roles in insulin-regulated glucose uptake. *Trends Endocrinol Metab* **11**, 356-361, (2000).
17 Holman, G. D. & Sandoval, I. V. Moving the insulin-regulated glucose transporter GLUT4 into and out of storage. *Trends Cell Biol* **11**, 173-179, (2001).
18 Bell, G. I., Burant, C. F., Takeda, J. & Gould, G. W. Structure and function of mammalian facilitative sugar transporters. *J Biol Chem* **268**, 19161-19164 (1993).
19 Agius, L. & Peak, M. Binding and translocation of glucokinase in hepatocytes. *Biochem Soc Trans* **25**, 145-150 (1997).
20 Lillie, S. H. & Pringle, J. R. Reserve carbohydrate metabolism in Saccharomyces cerevisiae: responses to nutrient limitation. *J Bacteriol* **143**, 1384-1394 (1980).
21 Francois, J. & Parrou, J. L. Reserve carbohydrates metabolism in the yeast Saccharomyces cerevisiae. *FEMS Microbiol Rev* **25**, 125-145, (2001).
22 Neves, M. J. & Francois, J. On the mechanism by which a heat shock induces trehalose accumulation in Saccharomyces cerevisiae. *Biochem J* **288 (Pt 3)**, 859-864 (1992).
23 Farkas, I., Hardy, T. A., DePaoli-Roach, A. A. & Roach, P. J. Isolation of the GSY1 gene encoding yeast glycogen synthase and evidence for the existence of a second gene. *J Biol Chem* **265**, 20879-20886 (1990).
24 Ni, H. T. & LaPorte, D. C. Response of a yeast glycogen synthase gene to stress. *Mol Microbiol* **16**, 1197-1205 (1995).
25 Thevelein, J. M. & de Winde, J. H. Novel sensing mechanisms and targets for the cAMP-protein kinase A pathway in the yeast Saccharomyces cerevisiae. *Mol Microbiol* **33**, 904-918, (1999).
26 Weinstein, D. A., Correia, C. E., Saunders, A. C. & Wolfsdorf, J. I. Hepatic glycogen synthase deficiency: an infrequently recognized cause of ketotic hypoglycemia. *Mol Genet Metab* **87**, 284-288, (2006).
27 Orho, M. *et al.* Mutations in the liver glycogen synthase gene in children with hypoglycemia due to glycogen storage disease type 0. *J Clin Invest* **102**, 507-515, (1998).
28 Soggia, A. P., Correa-Giannella, M. L., Fortes, M. A., Luna, A. M. & Pereira, M. A. A novel mutation in the glycogen synthase 2 gene in a child with glycogen storage disease type 0. *BMC Med Genet* **11**, 3, (2010).
29 Kollberg, G. *et al.* Cardiomyopathy and exercise intolerance in muscle glycogen storage disease 0. *N Engl J Med* **357**, 1507-1514, (2007).
30 McCue, M. E., Valberg, S. J., Lucio, M. & Mickelson, J. R. Glycogen synthase 1 (GYS1) mutation in diverse breeds with polysaccharide storage myopathy. *J Vet Intern Med* **22**, 1228-1233, (2008).
31 McCue, M. E. *et al.* Glycogen synthase (GYS1) mutation causes a novel skeletal muscle glycogenosis. *Genomics* **91**, 458-466, (2008).

32 Thomas, J. A., Schlender, K. K. & Larner, J. A rapid filter paper assay for UDPglucose-glycogen glucosyltransferase, including an improved biosynthesis of UDP-14C-glucose. *Anal Biochem* **25**, 486-499 (1968).
33 Hardy, T. A., Huang, D. & Roach, P. J. Interactions between cAMP-dependent and SNF1 protein kinases in the control of glycogen accumulation in Saccharomyces cerevisiae. *J Biol Chem* **269**, 27907-27913 (1994).
34 Guinovart, J. J. *et al.* Glycogen synthase: a new activity ratio assay expressing a high sensitivity to the phosphorylation state. *FEBS Lett* **106**, 284-288, (1979).
35 Withers, S. G. Structure and Mechanism of Glycosyl Transferases. *Annu Rev Biochem*, 75, (2006).
36 Lairson, L. L., Henrissat, B., Davies, G. J. & Withers, S. G. Glycosyltransferases: structures, functions, and mechanisms. *Annu Rev Biochem* **77**, 521-555, (2008).
37 Koshland, D. E. Stereochemistry and the mechanism of enzymatic reactions. *Biol. Rev. Camb. Phil. Soc* **28**, 416-436 (1953).
38 Cid, E., Gomis, R. R., Geremia, R. A., Guinovart, J. J. & Ferrer, J. C. Identification of two essential glutamic acid residues in glycogen synthase. *J Biol Chem* **275**, 33614-33621, (2000).
39 Phillips, D. C. The hen egg-white lysozyme molecule. *Proc. Natl. Acad. Sci. USA* **57**, 484–495, (1967).
40 Klein, H. W., Im, M. J. & Palm, D. Mechanism of the phosphorylase reaction. Utilization of D-gluco-hept-1-enitol in the absence of primer. *Eur J Biochem* **157**, 107-114 (1986).
41 Persson, K. *et al.* Crystal structure of the retaining galactosyltransferase LgtC from Neisseria meningitidis in complex with donor and acceptor sugar analogs. *Nat Struct Biol* **8**, 166-175, (2001).
42 Gibson, R. P., Turkenburg, J. P., Charnock, S. J., Lloyd, R. & Davies, G. J. Insights into trehalose synthesis provided by the structure of the retaining glucosyltransferase OtsA. *Chem Biol* **9**, 1337-1346, (2002).
43 Sommer, N., Depping, R., Piotrowski, M. & Ruger, W. Bacteriophage T4 alpha-glucosyltransferase: a novel interaction with gp45 and aspects of the catalytic mechanism. *Biochem Biophys Res Commun* **323**, 809-815, (2004).
44 Larner, J., Takeda, Y. & Hizukuri, S. The influence of chain size and molecular weight on the kinetic constants for the span glucose to polysaccharide for rabbit muscle glycogen synthase. *Mol Cell Biochem* **12**, 131-136 (1976).
45 Roach, P. J. Control of glycogen synthase by hierarchal protein phosphorylation. *FASEB J* **4**, 2961-2968 (1990).
46 Roach, P. J. Multisite and hierarchal protein phosphorylation. *J Biol Chem* **266**, 14139-14142 (1991).

47 Roach, P. J., Takeda, Y. & Larner, J. Rabbit skeletal muscle glycogen synthase. I. Relationship between phosphorylation state and kinetic properties. *J Biol Chem* **251**, 1913-1919 (1976).
48 Skurat, A. V., Wang, Y. & Roach, P. J. Rabbit skeletal muscle glycogen synthase expressed in COS cells. Identification of regulatory phosphorylation sites. *J Biol Chem* **269**, 25534-25542 (1994).
49 Skurat, A. V. & Roach, P. J. Phosphorylation of sites 3a and 3b (Ser640 and Ser644) in the control of rabbit muscle glycogen synthase. *J Biol Chem* **270**, 12491-12497 (1995).
50 Stralfors, P., Hiraga, A. & Cohen, P. The protein phosphatases involved in cellular regulation. Purification and characterisation of the glycogen-bound form of protein phosphatase-1 from rabbit skeletal muscle. *Eur J Biochem* **149**, 295-303 (1985).
51 Doherty, M. J., Moorhead, G., Morrice, N., Cohen, P. & Cohen, P. T. Amino acid sequence and expression of the hepatic glycogen-binding (GL)-subunit of protein phosphatase-1. *FEBS Lett* **375**, 294-298, (1995).
52 Doherty, M. J., Young, P. R. & Cohen, P. T. Amino acid sequence of a novel protein phosphatase 1 binding protein (R5) which is related to the liver- and muscle-specific glycogen binding subunits of protein phosphatase 1. *FEBS Lett* **399**, 339-343, (1996).
53 Armstrong, C. G., Browne, G. J., Cohen, P. & Cohen, P. T. PPP1R6, a novel member of the family of glycogen-targetting subunits of protein phosphatase 1. *FEBS Lett* **418**, 210-214, (1997).
54 Suzuki, Y. *et al.* Insulin control of glycogen metabolism in knockout mice lacking the muscle-specific protein phosphatase PP1G/RGL. *Mol Cell Biol* **21**, 2683-2694, (2001).
55 Doherty, M. J., Cadefau, J., Stalmans, W., Bollen, M. & Cohen, P. T. Loss of the hepatic glycogen-binding subunit (GL) of protein phosphatase 1 underlies deficient glycogen synthesis in insulin-dependent diabetic rats and in adrenalectomized starved rats. *Biochem J* **333 (Pt 2)**, 253-257 (1998).
56 Skurat, A. V., Dietrich, A. D. & Roach, P. J. Glycogen synthase sensitivity to insulin and glucose-6-phosphate is mediated by both NH2- and COOH-terminal phosphorylation sites. *Diabetes* **49**, 1096-1100 (2000).
57 Hojlund, K. *et al.* Dysregulation of glycogen synthase COOH- and NH2-terminal phosphorylation by insulin in obesity and type 2 diabetes mellitus. *J Clin Endocrinol Metab* **94**, 4547-4556, (2009).
58 Parker, G. J., Lund, K. C., Taylor, R. P. & McClain, D. A. Insulin resistance of glycogen synthase mediated by o-linked N-acetylglucosamine. *J Biol Chem* **278**, 10022-10027, (2003).
59 Hardy, T. A. & Roach, P. J. Control of yeast glycogen synthase-2 by COOH-terminal phosphorylation. *J Biol Chem* **268**, 23799-23805 (1993).

60 Huang, D., Farkas, I. & Roach, P. J. Pho85p, a cyclin-dependent protein kinase, and the Snf1p protein kinase act antagonistically to control glycogen accumulation in Saccharomyces cerevisiae. *Mol Cell Biol* **16**, 4357-4365 (1996).
61 Huang, D. *et al.* Cyclin partners determine Pho85 protein kinase substrate specificity in vitro and in vivo: control of glycogen biosynthesis by Pcl8 and Pcl10. *Mol Cell Biol* **18**, 3289-3299 (1998).
62 Pederson, B. A., Cheng, C., Wilson, W. A. & Roach, P. J. Regulation of glycogen synthase. Identification of residues involved in regulation by the allosteric ligand glucose-6-P and by phosphorylation. *J Biol Chem* **275**, 27753-27761, (2000).
63 Francois, J. M. *et al.* GAC1 may encode a regulatory subunit for protein phosphatase type 1 in Saccharomyces cerevisiae. *EMBO J* **11**, 87-96 (1992).
64 Stuart, J. S., Frederick, D. L., Varner, C. M. & Tatchell, K. The mutant type 1 protein phosphatase encoded by glc7-1 from Saccharomyces cerevisiae fails to interact productively with the GAC1-encoded regulatory subunit. *Mol Cell Biol* **14**, 896-905 (1994).
65 Roach, P. J. & Larner, J. Rabbit skeletal muscle glycogen synthase. II. Enzyme phosphorylation state and effector concentrations as interacting control parameters. *J Biol Chem* **251**, 1920-1925 (1976).
66 Hanashiro, I. & Roach, P. J. Mutations of muscle glycogen synthase that disable activation by glucose 6-phosphate. *Arch Biochem Biophys* **397**, 286-292, (2002).
67 Ferrer, J. C., Baque, S. & Guinovart, J. J. Muscle glycogen synthase translocates from the cell nucleus to the cystosol in response to glucose. *FEBS Lett* **415**, 249-252, (1997).
68 Ou, H., Yan, L., Osmanovic, S., Greenberg, C. C. & Brady, M. J. Spatial reorganization of glycogen synthase upon activation in 3T3-L1 adipocytes. *Endocrinology* **146**, 494-502, (2005).
69 Fernandez-Novell, J. M., Bellido, D., Vilaro, S. & Guinovart, J. J. Glucose induces the translocation of glycogen synthase to the cell cortex in rat hepatocytes. *Biochem J* **321 (Pt 1)**, 227-231 (1997).
70 Prats, C. *et al.* Phosphorylation-dependent translocation of glycogen synthase to a novel structure during glycogen resynthesis. *J Biol Chem* **280**, 23165-23172, (2005).
71 Coutinho, P. M., Deleury, E., Davies, G. J. & Henrissat, B. An evolving hierarchical family classification for glycosyltransferases. *J Mol Biol* **328**, 307-317, (2003).
72 Sheng, F., Jia, X., Yep, A., Preiss, J. & Geiger, J. H. The crystal structures of the open and catalytically competent closed conformation of Escherichia coli glycogen synthase. *J Biol Chem* **284**, 17796-17807, (2009).

73 Sheng, F., Yep, A., Feng, L., Preiss, J. & Geiger, J. H. Oligosaccharide binding in Escherichia coli glycogen synthase. *Biochemistry* **48**, 10089-10097, (2009).
74 Buschiazzo, A. *et al.* Crystal structure of glycogen synthase: homologous enzymes catalyze glycogen synthesis and degradation. *EMBO J* **23**, 3196-3205, (2004).
75 Horcajada, C., Guinovart, J. J., Fita, I. & Ferrer, J. C. Crystal structure of an archaeal glycogen synthase: insights into oligomerization and substrate binding of eukaryotic glycogen synthases. *J Biol Chem* **281**, 2923-2931, (2006).
76 D, H. The Application of the Three-Dimensional Patterson Method and the Crystal Structures of Proustite, Ag3AsS3, and Pyrargyrite, Ag3SbS3. *J. Chem. Physics* **4**, 381-390 (1936).
77 D, H. The Determination of the Phases of the Structure Factors of Non-Centrosymmetric Crystals by the Method of Double Isomorphous Replacement. *Acta Cryst.* **9**, 1-9 (1956).
78 Perutz, M. F. Isomorphous Replacement and Phase Determination in Non-centrosymmetric Space Groups. *Acta Cryst.* **9**, 867-873 (1956).
79 Blaschke, U. K., Silberstein, J. & Muir, T. W. Protein engineering by expressed protein ligation. *Methods Enzymol* **328**, 478-496 (2000).
80 Holford, M. & Muir, T. W. Adding 'splice' to protein engineering. *Structure* **6**, 951-956 (1998).
81 David, R., Richter, M. P. & Beck-Sickinger, A. G. Expressed protein ligation. Method and applications. *Eur J Biochem* **271**, 663-677,(2004).
82 Costa, G. L., Bauer, J. C., McGowan, B., Angert, M. & Weiner, M. P. Site-directed mutagenesis using a rapid PCR-based method. *Methods Mol Biol* **57**, 239-248, (1996).
83 Schneider, G. & Lindqvist, Y. Ta6Br14 is a useful cluster compound for isomorphous replacement in protein crystallography. *Acta Crystallogr D Biol Crystallogr* **50**, 186-191, (1994).
84 Minor, W., Tomchick, D. & Otwinowski, Z. Strategies for macromolecular synchrotron crystallography. *Structure* **8**, R105-110, (2000).
85 Lee, B. & Richards, F. M. The interpretation of protein structures: estimation of static accessibility. *J Mol Biol* **55**, 379-400 (1971).
86 Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774-797, (2007).
87 Terwilliger, T. SOLVE and RESOLVE: automated structure solution, density modification and model building. *J Synchrotron Radiat* **11**, 49-52, (2004).
88 Terwilliger, T. C. & Berendzen, J. Automated MAD and MIR structure solution. *Acta Crystallogr D Biol Crystallogr* **55**, 849-861, (1999).
89 The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* **50**, 760-763, (1994).

90 Adams, P. D. *et al.* PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr D Biol Crystallogr* **58**, 1948-1954, (2002).
91 Terwilliger, T. C. *et al.* Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr D Biol Crystallogr* **64**, 61-69, (2008).
92 Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132, (2004).
93 Vaguine, A. A., Richelle, J. & Wodak, S. J. SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. *Acta Crystallogr D Biol Crystallogr* **55**, 191-205, (1999).
94 Li, X., Romero, P., Rani, M., Dunker, A. K. & Obradovic, Z. Predicting Protein Disorder for N-, C-, and Internal Regions. *Genome Inform Ser Workshop Genome Inform* **10**, 30-40 (1999).
95 Zhang, W., DePaoli-Roach, A. A. & Roach, P. J. Mechanisms of multisite phosphorylation and inactivation of rabbit muscle glycogen synthase. *Arch Biochem Biophys* **304**, 219-225, (1993).
96 Camici, M., DePaoli-Roach, A. A. & Roach, P. J. Rabbit liver glycogen synthase. Purification and comparison of the properties of glucose-6-P-dependent and glucose-6-P-independent forms of the enzyme. *J Biol Chem* **259**, 3429-3434 (1984).

CURRICULUM VITAE

Sulochanadevi Baskaran

EDUCATION

2010	Indiana University, IN PhD in Biochemistry and Molecular Biology
2003	Madurai Kamaraj University, Madurai,India Advanced Diploma-Bioinformatics
2002	Madurai Kamaraj University, Madurai, India M.Sc Microbial Gene Technology
2000	TamilNadu Dr.MGR Medical University, Madras, India B. Pharmacy

RESEARCH / WORK EXPERIENCE

8/05-8/10	Graduate Student (Ph.D), Indiana University School of Medicine, Indianapolis, IN
7/04-6/05	Lecturer, Department of Biotechnology and Bioinformatics, SRM University, Madras, India
6/03-6/04	Research Assistant, Amrutanjan Health Care limited, Madras, India
1/03-5/03	Bioinformatics Student Research Assistant, Center for Biotechnology, Anna University, Madras, India
1/01-4/02	Teaching Assistant, Madurai Kamaraj University, Madurai, India
6/01-5/02	Graduate Research Student (Masters) Madurai Kamaraj University, Madurai,India
10/99-4/00	Undergraduate Research Assistant, SRM College of Pharmacy, Madras, India
8/98-2/99	Hospital and Clinical Pharmacy Intern, Government Royapettah Hospital, Madras, India

AWARDS AND SCHOLARSHIPS

2010	Honorary mention, Sigma-Xi Research Competition
2010	Best Poster Award, Biochemistry Research Day
2008	Jack Davis Award for best seminar by a graduate student
2002-2003	Bioinformatics Student fellowship
1994-2000	National Merit Scholarship

2000	Best Outgoing Pharmacy Student
1995	Best Student (High School) of the District of Central Madras

PROFESSIONAL ACTIVITIES

Professional membership - American Crystallographic Association
Registered pharmacist, Pharmacy Council of India (2000-2005)

PRESENTATIONS

Baskaran S, Roach PJ, DePaoli-Roach AA, Hurley TD. Three state structural model for regulation of yeast glycogen synthase. Poster presentation, Biochemistry Research Day 2010, Indiana University School of Medicine

Baskaran S, Roach PJ, DePaoli-Roach AA, Hurley TD. Crystal structure of yeast glycogen synthase. Poster presentation, NIH-National Graduate Research Festival 2009, Bethesda

Baskaran S, Roach PJ, Hurley TD. Crystal structure of yeast glycogen synthase. Indiana Clinical and Translational Sciences Institute, Annual poster session 2009

Baskaran S, Roach PJ, Hurley TD. Structural and functional studies of yeast glycogen synthase, Poster presentation, ACA meet 2009, Toronto

PUBLICATIONS

Baskaran S, Roach PJ, DePaoli-Roach AA, Hurley TD. Structural basis for glucose-6-phosphate activation of glycogen synthase. (Submitted to PNAS)

Baskaran S, Roach PJ, DePaoli-Roach AA, Hurley TD. Conserved glycogen binding sites on the surface of glycogen synthase are required for efficient catalysis. (Manuscript in preparation)

Travers JB, Edenberg HJ, Zhang Q, Al-Hassani M, Yi Q, Baskaran S, Konger RL. Augmentation of UVB radiation-mediated early gene expression by the epidermal platelet-activating factor receptor. J Invest Dermatol. 2008;128(2):455-60.